

# FREE CONVEX ALGEBRAIC GEOMETRY

J. WILLIAM HELTON<sup>1</sup>, IGOR KLEP<sup>2</sup>, AND SCOTT MCCULLOUGH<sup>3</sup>

**ABSTRACT.** This chapter is a tutorial on techniques and results in free convex algebraic geometry and free real algebraic geometry (RAG). The term free refers to the central role played by algebras of noncommutative polynomials  $\mathbb{R}\langle x \rangle$  in free (freely noncommuting) variables  $x = (x_1, \dots, x_g)$ . The subject pertains to problems where the unknowns are matrices or Hilbert space operators as arise in linear systems engineering and quantum information theory.

The subject of free RAG flows in two branches. One, free positivity and inequalities is an analog of classical real algebraic geometry, a theory of polynomial inequalities embodied in algebraic formulas called Positivstellensätze; often free Positivstellensätze have cleaner statements than their commutative counterparts. Free convexity, the second branch of free RAG, arose in an effort to unify a torrent of ad hoc optimization techniques which came on the linear systems engineering scene in the mid 1990's. Mathematically, much as in the commutative case, free convexity is connected with free positivity through the second derivative: A free polynomial is convex if and only if its Hessian is positive. However, free convexity is a very restrictive condition, for example, free convex polynomials have degree 2 or less.

This article describes for a beginner techniques involving free convexity. As such it also serves as a point of entry into the larger field of free real algebraic geometry.

## 1. INTRODUCTION

This chapter is a tutorial on techniques and results in *free convex algebraic geometry* and *free positivity*. As such it also serves as a point of entry into the larger field of *free real algebraic geometry* (*free RAG*), and makes contact with noncommutative real algebraic geometry [Hel02, HKM10c, HKM13, HKM12a, HM12, KS08a,

---

*Date:* April 17, 2013.

2000 *Mathematics Subject Classification.* 47A63, 46L89, 14P10 (Primary), 15A22, 13J30 (Secondary).

*Key words and phrases.* Noncommutative polynomial, Linear Matrix Inequality, convexity, positivity, rational function, middle matrix, free positivity, free convexity.

<sup>1</sup>Research supported by NSF grants DMS-0700758, DMS-0757212, and the Ford Motor Co.

<sup>2</sup>Supported by the Faculty Research Development Fund (FRDF) of The University of Auckland (project no. 3701119). Partially supported by the Slovenian Research Agency grant P1-0222.

<sup>3</sup>Research supported by the NSF grants DMS-0758306 and DMS-1101137.

[KS08b](#), [McC01](#), [PNA10](#), [Smü05](#), [Smü09](#)], free analysis and free probability (lying at the origins of free analysis, cf. [\[SV06\]](#)), free analytic function theory and free harmonic analysis [[HKM10a](#), [HKM10b](#), [HKMS09](#), [MS11](#), [Pop06](#), [Voi04](#), [Voi10](#), [KVV+](#)].

The term free here refers to the central role played by algebras of noncommutative polynomials  $\mathbb{R}\langle x \rangle$  in free (freely noncommuting) variables  $x = (x_1, \dots, x_g)$ . A striking difference between the free and classical settings is the following Positivstellensatz.

**Theorem 1** (Helton [\[Hel02\]](#)). *A nonnegative (suitably defined) free polynomial is a sum of squares.*

The subject of free RAG flows in two branches. One, free positivity is an analog of classical real algebraic geometry, a theory of polynomial inequalities embodied in Positivstellensätze. As is the case with the sum of squares result above (Theorem 1), generally free Positivstellensätze have cleaner statements than do their commutative counterparts; see e.g. [\[McC01, Hel02, HMP04, HKM12a\]](#) for a sample. Free convexity, the second branch of free RAG, arose in an effort to unify a torrent of ad hoc techniques which came on the linear systems engineering scene in the mid 1990's. We soon give a quick sketch of the engineering motivation, based on the slightly more complete sketch given in the survey article [\[dOHMP09\]](#). Mathematically, much as in the commutative case, free convexity is connected with free positivity through the second derivative: A free polynomial is convex if and only if its Hessian is positive.

The tutorial proper starts with Section 2. In the remainder of this introduction, motivation for the study of free positivity and convexity arising in *linear systems engineering*, *quantum phenomena*, and other subjects such as *free probability* is provided, as are some suggestions for further reading.

**1.1. Motivation.** While the theory is both mathematically pleasing and natural, much of the excitement of free convexity and positivity stems from its applications. Indeed, the fact that a large class of linear systems engineering problems naturally lead to free inequalities provided the main force behind the development of the subject. In this motivational section, we describe in some detail the linear systems point of view. We also give a brief introduction to other applications.

**1.1.1. Linear Systems Engineering.** The layout of a linear systems problem is typically specified by a signal flow diagram. Signals go into boxes and other signals come out. The boxes in a linear system contain constant coefficient linear differential equations which are specified entirely by matrices (the coefficients of the differential equations). Often many boxes appear and many signals transmit between them. In a

typical problem some boxes are given and some we get to design subject to the condition that the  $L^2$  norm of various signals must compare in a prescribed way, e.g. the input to the system has  $L^2$  norm bigger than the output. The signal flow diagram itself and corresponding problems do not specify the size of matrices involved. So ideally any algorithms derived apply to matrices of all sizes. Hence the problems are called *dimension free*.

An empirical observation is that system problems of this type convert to inequalities on polynomials in matrices, the form of the polynomials being determined entirely by the signal flow layout (and independent of the matrices involved). Thus the systems problem naturally leads to free polynomials and free positivity conditions.

For yet a more detailed discussion of this example, see [dOHMP09, §4.1]. Those who read Chapter 2 saw a basic example of this in Chapter 2.2.1. Next we give more of an idea of how the correspondence between linear systems and noncommutative polynomials occurs. This is done primarily with an example.

1.1.2. *Linear systems.* A *linear system*  $\mathfrak{F}$  is given by the constant coefficient linear differential equations

$$\begin{aligned}\frac{dx}{dt} &= Ax + Bu, \\ y &= Cx,\end{aligned}$$

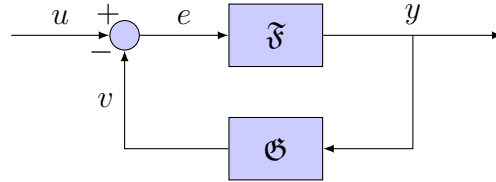
with the vector

- $x(t)$  at each time  $t$  being in the vector space  $\mathcal{X}$  called the *state space*,
- $u(t)$  at each time  $t$  being in the vector space  $\mathcal{U}$  called the *input space*,
- $y(t)$  at each time  $t$  being in the vector space  $\mathcal{Y}$  called the *output space*,

and  $A, B, C$  being linear maps on the corresponding vector spaces.

1.1.3. *Connecting linear systems.* Systems can be connected in incredibly complicated configurations. We describe a simple connection and this goes a long way toward illustrating the general idea. Given two linear systems  $\mathfrak{F}$ ,  $\mathfrak{G}$ , we describe the formulas for connecting them in feedback.

One basic feedback connection is described by the diagram



called a *signal flow diagram*. Here  $u$  is a signal going into the *closed loop system* and  $y$  is the signal coming out. The signal flow diagram is equivalent to a collection of equations. The systems  $\mathfrak{F}$  and  $\mathfrak{G}$  themselves are respectively given by the linear differential equations

$$\begin{aligned} \frac{dx}{dt} &= Ax + Be, & \frac{d\xi}{dt} &= Q\xi + R w, \\ y &= Cx, & v &= S\xi. \end{aligned}$$

The feedback connection is described algebraically by

$$w = y \quad \text{and} \quad e = u - v.$$

Putting these relations together gives that the closed loop system is described by differential equations

$$\begin{aligned} \frac{dx}{dt} &= Ax - BS\xi + Bu, \\ \frac{d\xi}{dt} &= Q\xi + Ry = Q\xi + RCx, \\ y &= Cx. \end{aligned}$$

which is conveniently described in matrix form as

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} x \\ \xi \end{bmatrix} &= \begin{bmatrix} A & -BS \\ RC & Q \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u, \\ y &= [C \ 0] \begin{bmatrix} x \\ \xi \end{bmatrix}, \end{aligned} \tag{1}$$

where the state space of the closed loop systems is the direct sum  $\mathcal{X} \oplus \mathcal{Y}$  of the state spaces  $\mathcal{X}$  of  $\mathfrak{F}$  and  $\mathcal{Y}$  of  $\mathfrak{G}$ . From (1), the coefficients of the O.D.E. are (block) matrices whose entries are (in this case simple) polynomials in the matrices  $A, B, C, Q, R, S$ .

This illustrates the moral of the general story:

*System connections produce a new system whose coefficients are matrices with entries which are noncommutative polynomials (or at worst “rational expressions”) in the coefficient matrices of the component systems.*

Complicated signal flow diagrams give complicated matrices of noncommutative polynomials or rationals. Note in what was said the dimensions of vector spaces and

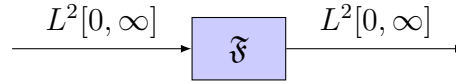
matrices  $A, B, C, Q, R, S$  never entered explicitly; the algebraic form of (1) is completely determined by the flow diagram. Thus, such linear systems lead to *dimension free* problems.

Next we turn to how “noncommutative inequalities” arise. The main constraint producing them can be thought of as energy dissipation, a special case of which are the Lyapunov functions already seen in Chapter 2.2.1.

1.1.4. *Energy dissipation.* We have a system  $\mathfrak{F}$  and want a condition which checks whether

$$\int_0^\infty |u|^2 dt \geq \int_0^\infty |\mathfrak{F}u|^2 dt, \quad x(0) = 0,$$

holds for all input functions  $u$ , where  $\mathfrak{F}u = y$  in the above notation. If this holds  $\mathfrak{F}$  is called a *dissipative system*.



The energy dissipative condition is formulated in the language of analysis, but it converts to algebra (or at least an algebraic inequality) because of the following construction, which assumes the existence of a “potential energy”-like function  $V$  on the state space. A function  $V$  which satisfies  $V \geq 0$ ,  $V(0) = 0$ , and

$$V(x(t_1)) + \int_{t_1}^{t_2} |u(t)|^2 dt \geq V(x(t_2)) + \int_{t_1}^{t_2} |y(t)|^2 dt$$

for all input functions  $u$  and initial states  $x_1$  is called a *storage function*. The displayed inequality is interpreted physically as

$$\text{potential energy now} + \text{energy in} \geq \text{potential energy then} + \text{energy out}.$$

Assuming enough smoothness of  $V$ , we can differentiate this integral condition and use  $\frac{d}{dt}x(t_1) = Ax(t_1) + Bu(t_1)$  to obtain a differential inequality

$$0 \geq \nabla V(x)(Ax + Bu) + |Cx|^2 - |u|^2, \quad (2)$$

on what is called the “reachable set” (which we do not need to define here).

In the case of linear systems,  $V$  can be chosen to be a quadratic. So it has the form  $V(x) = \langle Ex, x \rangle$  with  $E \succeq 0$  and  $\nabla V(x) = 2Ex$ .

**Theorem 2.** *The linear system  $A, B, C$  is dissipative if inequality (2) holds for all  $u \in \mathcal{U}, x \in \mathcal{X}$ . Conversely, if  $A, B, C$  is “reachable”<sup>1</sup>, then dissipativity implies inequality (2) holds for all  $u \in \mathcal{U}, x \in \mathcal{X}$ .*

In the linear case, we may substitute  $\nabla V(x) = 2Ex$  in (2) to obtain

$$0 \geq 2(Ex)^\top(Ax + Bu) + |Cx|^2 - |u|^2,$$

for all  $u, x$ . Then maximize in  $x$  to get

$$0 \geq x^\top[EA + A^\top E + EBB^\top E + C^\top C]x.$$

Thus the classical *Riccati matrix inequality*

$$0 \succeq EA + A^\top E + EBB^\top E + C^\top C \quad \text{with} \quad E \succeq 0 \quad (3)$$

ensures dissipativity of the system; and, it turns out, is also implied by dissipativity when the system is reachable.

It is inequality (3), applied in many many contexts, which leads to positive semi-definite inequalities throughout all of linear systems theory.

As an aside we return to the very special case of dissipativity, namely Lyapunov stability, described in Chapter 2.2.1. Our discussion starts with the “miracle of inequality (3)”: when  $B = 0$  it becomes the Lyapunov inequality. However, this is merely magic (no miracle whatsoever); the trick being that the if input  $u$  is identically zero, then dissipativity implies stability. The converse is less intuitive, but true: stability of  $\dot{x} = Ax$  implies existence of a “virtual” potential energy  $V(x) = \langle Ex, x \rangle$  and output  $C$  making the “virtual” system dissipative.

**1.1.5. Schur Complements and Linear Matrix Inequalities.** Using Schur complements, the Riccati inequality of equation (3) is equivalent to the inequality

$$L(E) := \begin{bmatrix} EA + A^\top E + C^\top C & EB \\ B^\top E & -I \end{bmatrix} \preceq 0.$$

Here  $A, B, C$  describe the system and  $E$  is an unknown matrix. If the system is reachable, then  $A, B, C$  is dissipative if and only if  $L(E) \preceq 0$  and  $E \succeq 0$ .

The key feature in this reformulation of the Riccati inequality is that  $L(E)$  is linear in  $E$ , so the inequality  $L(E) \preceq 0$  is a *Linear Matrix Inequality* (LMI) in  $E$ .

---

<sup>1</sup>A mild technical condition

1.1.6. *Putting it together.* We have shown two ingredients of linear system theory, connection laws (algebraic) and dissipation (inequalities), but have yet to put them together. It is in fact a very mechanical procedure. After going through the procedure one sees that the problem a software toolbox designer faces is this:

(GRAIL) Given a symmetric matrix of nc polynomials

$$p(a, x) = \left[ p_{ij}(a, x) \right]_{i,j=1}^k,$$

and a tuple of matrices  $A$ , provide an algorithm for finding  $X$  making  $p(A, X) \succeq 0$  or better yet as large as possible.

Algorithms for doing this are based on numerical optimization or a close relative, so even if they find a local solution there is no guarantee that it is global. If  $p$  is convex in  $X$ , then these problems disappear.

Thus, systems problems described by signal flow diagrams produce a mess of matrix inequalities with some matrices known and some unknown and the constraints that some polynomials are positive semidefinite. The inequalities can get very complicated as one might guess, since signal flow diagrams get complicated. These considerations thus naturally lead to the emerging subject of free real algebraic geometry, the study of noncommutative (free) polynomial inequalities and free semialgebraic sets. Indeed, much of what is known about this very new subject is touched on in this chapter.

The engineer would like for these polynomial inequalities to be convex in the unknowns. Convexity guarantees that local optima are global optima (finding global optima is often of paramount importance) and facilitates numerics.

Hence the major issues in linear systems theory are:

- (1) *Which problems convert to a convex matrix inequality? How does one do the conversion?*
- (2) *Find numerics which will solve large convex problems. How do you use special structure, such as most unknowns are matrices and the formulas are all built of noncommutative rational functions?*
- (3) *Are convex matrix inequalities more general than LMIs?*

The mathematics here can be motivated by the problem of writing a toolbox for engineers to use in designing linear systems. What goes in such toolboxes is algebraic formulas with matrices  $A, B, C$  unspecified and reliable numerics for solving them when a user does specify  $A, B, C$  as matrices. A user who designs a controller for a helicopter puts in the mathematical systems model for his helicopter and puts in

matrices, for example,  $A$  is a particular  $8 \times 8$  real matrix etc. Another user who designs a satellite controller might have a 50 dimensional state space and of course would pick completely different  $A, B, C$ . Essentially any matrices of any compatible dimensions can occur. Any claim we make about our formulas must be valid regardless of the size of the matrices plugged in.

The toolbox designer faces two completely different tasks. One is manipulation of algebraic inequalities; the other is numerical solutions. Often the first is far more daunting since the numerics is handled by some standard package (although for numerics problem size is a demon). Thus there is a great need for algebraic theory. Most of this chapter bears on questions like (3) above where the unknowns are matrices. *The first two questions will not be addressed.* Here we treat (3) when there are no  $a$  variables. When there are  $a$  variables see [HHLM08, BM+]. Thus we shall consider polynomials  $p(x)$  in free noncommutative variables  $x$  and focus on their convexity on free semialgebraic sets.

What are the implications of our study for engineering? Herein you will see strong results on free convexity but what do they say to an engineer? We foreshadow the forthcoming answer by saying it is fairly negative, but postpone further disclosure till the final page of these writings not so much to promote suspense, but for the conclusion to arrive after you have absorbed the theory.

1.1.7. *Quantum Phenomena.* Free Positivstellensätze - algebraic certificates for positivity - of which Theorem 1 is the grandad, have physical applications. Applications to quantum physics are explained by Pironio, Navascués, Acín [PNA10] who also consider computational aspects related to noncommutative sum of squares. How this pertains to operator algebras is discussed by Schweighofer and the second author in [KS08a]. The important Bessis-Moussa-Villani conjecture (BMV) from quantum statistical mechanics is tackled in [KS08b, CKP10]. Doherty, Liang, Toner, Wehner [DLTW08] employ noncommutative positivity and the Positivstellensatz [HM04b] of the first and the third author to consider the quantum moment problem and multiprover games.

A particularly elegant recent development, independent of the line of history containing the work in this chapter, was initiated by Effros. The classic “perspective” transformation carries a function on  $\mathbb{R}^n$  to a function on  $\mathbb{R}^{n+1}$ . It is used for various purposes, one being in algebraic geometry to produce “blowups” of singularities thereby removing them. It has the property that convex functions map to convex functions. What about convex functions on free variables? This question was asked by Effros and settled affirmatively in [Eff09] for natural cases as a way to show



that quantum relative entropy is convex. Subsequently, [ENG11] showed that the perspective transformation in free variables always maps convex functions to convex functions.

1.1.8. *Miscellaneous applications.* A number of other scientific disciplines use free analysis, though less systematically than in free real algebraic geometry.

**Free probability.** Voiculescu developed it to attack one of the purest of mathematical questions regarding von Neumann algebras. From the outset (about 20 years ago) it was elegant and it came to have great depth. Subsequently, it was discovered to bear forcefully and effectively on random matrices. The area is vast, so we do not dive in but refer the reader to an introduction [SV06, VDN92].

**Nonlinear engineering systems.** A classical technique in nonlinear systems theory developed by Fliess is based on manipulation of power series with noncommutative variables (the Chen series). The area has a new impetus coming from the problem of data compression, so now is a time when these correspondences are being worked out, cf. [GL05, GT12, LCL04].

1.2. **Further reading.** We pause here to offer some suggestions for further reading. For further engineering motivation we recommend the paper [SI95] or the longer version [SIG97] for related new directions. Descriptions of Positivstellensätze are in the surveys [HKM12b, dOHMP09, HP07, Smü09] with the first three also briskly touring free convexity. The survey article [HMPV09] is aimed at engineers.

Noncommutative is a broad term, encompassing essentially all algebras. In between the extremes of commutative and free lie many important topics, such as Lie algebras, Hopf algebras, quantum groups,  $C^*$ -algebras, von Neumann algebras, etc. For instance, there are elegant noncommutative real algebraic geometry results for the Weyl Algebra [Smü05], cf. [Smü09].

1.3. **Guide to the chapter.** The goal of this tutorial is to introduce the reader to the main results and techniques used to study free convexity. Fortunately, the subject is new and the techniques not too numerous so that one can quickly become an expert.

The basics of free, or nc, polynomials and their evaluations are developed in Section 2. The key notions are positivity and convexity for free polynomials. The principal fact is that the second directional derivative (in direction  $h$ ) of a free convex polynomial is a positive quadratic polynomial in  $h$  (just like in the commutative case). Free quadratic (in  $h$ ) polynomials have a Gram type representation which thus figures

prominently in studying convexity. The nuts and bolts of this Gram representation and some of its consequences, including Theorem 1, are the subjects of Sections 4 and 5 respectively.

The Gram representation techniques actually require only a small amount of convexity and thus there is a theory of geometry on free varieties having signed (e.g. positive) curvature. Some details are in Section 6.

A couple of free real algebraic geometry results which have a heavy convexity component are described in the last section, Section 7. The first is an optimal free convex Positivstellensatz which generalizes Theorem 1. The second says that free convex semialgebraic sets are free spectrahedra, giving another example of the much more rigid structure in the free setting.

Section 3 introduces software which handles free noncommutative computations. You may find it useful in your free studies.

In what follows, mildly incorrectly, but in keeping with the usage in the literature, the terms noncommutative (abbreviated nc) and free are used synonymously.

## 2. BASICS OF NC POLYNOMIALS AND THEIR CONVEXITY

This section treats the basics of polynomials in nc variables, nc differential calculus, and nc inequalities. There is also a brief introduction to nc rational functions and inequalities.

**2.1. Noncommutative polynomials.** Before turning to the formalities, we give, by examples, an informal introduction to noncommutative (nc) polynomials.

A noncommutative polynomial  $p$  is a polynomial in a finite set  $x = (x_1, \dots, x_g)$  of relation free variables. A canonical example, in the case of two variables  $x = (x_1, x_2)$ , is the *commutator*

$$c(x_1, x_2) = x_1x_2 - x_2x_1. \quad (4)$$

It is precisely the fact that  $x_1$  and  $x_2$  do not commute that makes  $c$  nonzero.

While a commutative polynomial  $q \in \mathbb{R}[t_1, t_2]$  is naturally evaluated at points  $t \in \mathbb{R}^2$ , nc polynomials are naturally evaluated on tuples of square matrices. For instance, with

$$X_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

and  $X = (X_1, X_2)$ , one finds

$$c(X) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Importantly,  $c$  can be evaluated on any pair  $(X, Y)$  of symmetric matrices of the same size. (Later in the section we will also consider evaluations involving not necessarily symmetric matrices.) Note that if  $X$  and  $Y$  are  $n \times n$ , then  $c(X, Y)$  is itself an  $n \times n$  matrix. In the case of  $c(x, y) = xy - yx$ , the matrix  $c(X, Y) = 0$  if and only if  $X$  and  $Y$  commute. In particular,  $c$  is zero on  $\mathbb{R}^2$  (2-tuples of  $1 \times 1$  matrices).

For another example, if  $d(x_1, x_2) = 1 + x_1 x_2 x_1$ , then with  $X_1$  and  $X_2$  as above, we find

$$d(X) = I_2 + X_1 X_2 X_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Note that although  $X$  is a tuple of symmetric matrices, it need not be the case that  $p(X)$  is symmetric. Indeed, the matrix  $c(X)$  above is not. In the present context, we say that  $p$  is *symmetric*, if  $p(X)$  is symmetric whenever  $X = (X_1, \dots, X_g)$  is a tuple of symmetric matrices. Another more algebraic definition of symmetric for nc polynomials appears in Section 2.2.

**2.1.1. Noncommutative convexity for polynomials.** Many standard notions for polynomials, and even functions, on  $\mathbb{R}^g$  extend to the nc setting, though often with unexpected ramifications. For example, the commutative polynomial  $q \in \mathbb{R}[t_1, t_2]$  is convex if, given  $s, t \in \mathbb{R}^2$ ,

$$\frac{1}{2}(q(s) + q(t)) \geq q\left(\frac{s+t}{2}\right).$$

There is a natural ordering on symmetric  $n \times n$  matrices defined by  $X \succeq Y$  if the symmetric matrix  $X - Y$  is positive semidefinite; i.e., if its eigenvalues are all nonnegative. Similarly,  $X \succ Y$ , if  $X - Y$  is positive definite; i.e., all its eigenvalues are positive. This order yields a canonical notion of convex nc polynomial. Namely, a symmetric polynomial  $p$  is *convex* if for each  $n$  and each pair of  $g$  tuples of  $n \times n$  symmetric matrices  $X = (X_1, \dots, X_g)$  and  $Y = (Y_1, \dots, Y_g)$ , we have

$$\frac{1}{2}(p(X) + p(Y)) \succeq p\left(\frac{X+Y}{2}\right).$$

Equivalently,

$$\frac{p(X) + p(Y)}{2} - p\left(\frac{X+Y}{2}\right) \succeq 0. \tag{5}$$

Even in one variable, convexity for an nc polynomial is a serious constraint. For instance, consider the polynomial  $x^4$ . It is symmetric, but with

$$X = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \text{ and } Y = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

it follows that

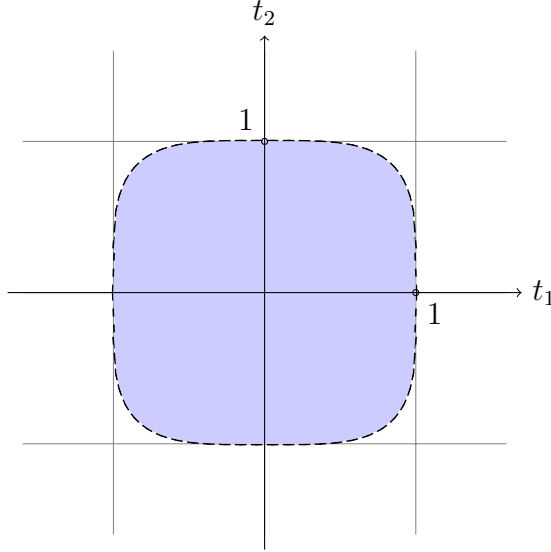
$$\frac{X^4 + Y^4}{2} - \left( \frac{1}{2}X + \frac{1}{2}Y \right)^4 = \begin{bmatrix} 164 & 120 \\ 120 & 84 \end{bmatrix}$$

is not positive semidefinite. Thus  $x^4$  is not convex.

**2.1.2. Noncommutative polynomial inequalities and convexity.** The study of polynomial inequalities, real algebraic geometry or semialgebraic geometry, has a nc version. A *basic open semialgebraic set* is a subset of  $\mathbb{R}^g$  defined by a list of polynomial inequalities; i.e., a set  $S$  is a basic open semialgebraic set if

$$S = \{t \in \mathbb{R}^g : p_1(t) > 0, \dots, p_k(t) > 0\}$$

for some polynomials  $p_1, \dots, p_k \in \mathbb{R}[t_1, \dots, t_g]$ .



$$\text{ncTV}(1) = \{(t_1, t_2) \in \mathbb{R}^2 : 1 - t_1^4 - t_2^4 > 0\}.$$

Because noncommutative polynomials are evaluated on tuples of matrices, a nc (free) basic open semialgebraic set is a sequence. For positive integers  $n$ , let  $(\mathbb{S}^{n \times n})^g$  denote the set of  $g$ -tuples of  $n \times n$  symmetric matrices. Given symmetric nc polynomials  $p_1, \dots, p_k$ , let

$$\mathcal{P}(n) = \{X \in (\mathbb{S}^{n \times n})^g : p_1(X) \succ 0, \dots, p_k(X) \succ 0\}.$$

The sequence  $\mathcal{P} = (\mathcal{P}(n))$  is then a *nc (free) basic open semialgebraic set*. The sequence

$$\text{ncTV}(n) = \{X \in (\mathbb{S}^{n \times n})^2 : I_n - X_1^4 - X_2^4 \succ 0\}$$

is an entertaining example. When  $n = 1$ ,  $\text{ncTV}(1)$  is a subset of  $\mathbb{R}^2$  often called the *TV screen*. Numerically it can be verified, though it rather tricky to do so (see Exercise 23) that the set  $\text{ncTV}(2)$  is not a convex set. An analytic proof that  $\text{ncTV}(n)$  is not a convex set for some  $n$  can be found in [DHM07a]. It also follows by combining results in [HM12] and [HV07]. For properties of the classical commutative TV screen, see the Chapters 6 of Nie and 5 by Rostalski-Sturmfels in this book.

**Example 3.** Let  $p_\epsilon := \epsilon^2 - \sum_{j=1}^g x_j^2$ . Then the  $\epsilon$ -neighborhood of 0,

$$\mathcal{N}_\epsilon := \bigcup_{n \in \mathbb{N}} \{X \in (\mathbb{S}^{n \times n})^g : p_\epsilon(X) \succ 0\}$$

is an important example of a nc basic open semialgebraic set.

**2.2. Noncommutative polynomials, the formalities.** We now take up the formalities of nc polynomials, their evaluations, convexity, and positivity.

Let  $x = (x_1, \dots, x_g)$  denote a  $g$ -tuple of free noncommuting variables and let  $\mathbb{R}\langle x \rangle$  denote the associative  $\mathbb{R}$ -algebra freely generated by  $x$ , i.e., the elements of  $\mathbb{R}\langle x \rangle$  are polynomials in the noncommuting variables  $x$  with coefficients in  $\mathbb{R}$ . Its elements are called (*nc*) *polynomials*. An element of the form  $aw$  where  $0 \neq a \in \mathbb{R}$  and  $w$  is a *word* in the variables  $x$  is called a *monomial* and  $a$  its *coefficient*. Hence words are monomials whose coefficient is 1. Note that the empty word  $\emptyset$  plays the role of the multiplicative identity for  $\mathbb{R}\langle x \rangle$ .

There is a natural *involution*  $^\top$  on  $\mathbb{R}\langle x \rangle$  that reverses words. For example,  $(2 - 3x_1^2 x_2 x_3)^\top = 2 - 3x_3 x_2 x_1^2$ . A polynomial  $p$  is a *symmetric polynomial* if  $p^\top = p$ . Later we will see that this notion of symmetric is equivalent to that in the previous subsection. For now we note that of

$$\begin{aligned} c(x) &= x_1 x_2 - x_2 x_1 \\ j(x) &= x_1 x_2 + x_2 x_1 \end{aligned}$$

$j$  is symmetric, but  $c$  is not. Indeed,  $c^\top = -c$ . Because  $x_j^\top = x_j$  we refer to the variables as *symmetric variables*. Occasionally we emphasize this point by writing  $\mathbb{R}\langle x = x^\top \rangle$  for  $\mathbb{R}\langle x \rangle$ .

The *degree* of an nc polynomial  $p$ , denoted  $\deg(p)$ , is the length of the longest word appearing in  $p$ . For instance the polynomials  $c$  and  $j$  above both have degree

two and the degree of

$$r(x) = 1 - 3x_1x_2 - 3x_2x_1 - 2x_1^2x_2^4x_1^2$$

is eight. Let  $\mathbb{R}\langle x \rangle_k$  denote the polynomials of degree at most  $k$ .

**2.2.1. Noncommutative matrix polynomials.** Given positive integers  $d, d' \in \mathbb{N}$ , let  $\mathbb{R}^{d \times d'} \langle x \rangle$  denote the  $d \times d'$  matrices with entries from  $\mathbb{R}\langle x \rangle$ . Thus elements of  $\mathbb{R}^{d \times d'} \langle x \rangle$  are *matrix-valued nc polynomials*. The involution on  $\mathbb{R}\langle x \rangle$  naturally extends to a mapping  $\top : \mathbb{R}^{d \times d'} \langle x \rangle \rightarrow \mathbb{R}^{d' \times d} \langle x \rangle$ . In particular, if

$$P = [p_{i,j}]_{i,j=1}^{d,d'} \in \mathbb{R}^{d \times d'} \langle x \rangle,$$

then

$$P^\top = [p_{j,i}^\top]_{i,j=1}^{d',d} \in \mathbb{R}^{d' \times d} \langle x \rangle.$$

In the case that  $d = d'$ , such a  $P$  is symmetric if  $P^\top = P$ .

**2.2.2. Linear pencils.** Given a positive integer  $n$ , let  $\mathbb{S}^{n \times n}$  denote the real symmetric  $n \times n$  matrices. For  $A_0, A_1, \dots, A_g \in \mathbb{S}^{d \times d}$ , the expression

$$L(x) = A_0 + \sum_{j=1}^g A_j x_j \in \mathbb{S}^{d \times d} \langle x \rangle \quad (6)$$

in the noncommuting variables  $x$  is a *symmetric affine linear pencil*. In other words, these are precisely the symmetric degree one matrix-valued nc polynomials. If  $A_0 = I$ , then  $L$  is *monic*. If  $A_0 = 0$ , then  $L$  is a *linear pencil*. The homogeneous linear part  $\sum_{j=1}^g A_j x_j$  of a linear pencil  $L$  as in (6) will be denoted by  $L^{(1)}$ .

**Example 4.** Let

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Then

$$I + \sum A_j x_j = \begin{bmatrix} 1 & x_1 & 0 & 0 \\ x_1 & 1 & x_2 & 0 \\ 0 & x_2 & 1 & x_3 \\ 0 & 0 & x_3 & 1 \end{bmatrix}$$

is the corresponding monic affine linear pencil.

**2.2.3. Polynomial evaluations.** If  $p \in \mathbb{R}^{d \times d'} \langle x \rangle$  is an nc polynomial and  $X \in (\mathbb{S}^{n \times n})^g$ , the *evaluation*  $p(X) \in \mathbb{R}^{dn \times d'n}$  is defined by simply replacing  $x_i$  by  $X_i$ . Throughout we use lower case letters for variables and the corresponding capital letter for matrices substituted for that variable.

**Example 5.** Suppose  $p(x) = Ax_1x_2$  where  $A = \begin{bmatrix} -4 & 2 \\ 3 & 0 \end{bmatrix}$ . That is,

$$p(x) = \begin{bmatrix} -4x_1x_2 & 2x_1x_2 \\ 3x_1x_2 & 0 \end{bmatrix}.$$

Thus  $p \in \mathbb{R}^{2 \times 2} \langle x \rangle$  and one example of an evaluation is

$$\begin{aligned} p\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) &= A \otimes \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) = A \otimes \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right) \\ &= \begin{bmatrix} 0 & 4 & 0 & -2 \\ -4 & 0 & 2 & 0 \\ 0 & -3 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Similarly, if  $p$  is a constant matrix-valued nc polynomial,  $p(x) = A$ , and  $X \in (\mathbb{S}^{n \times n})^g$ , then  $p(X) = A \otimes I_n$ . Here we have taken advantage of the usual *tensor* (or *Kronecker*) *product* of matrices. Given an  $\ell \times \ell'$  matrix  $A = (A_{i,j})$  and an  $n \times n'$  matrix  $B$ , by definition,  $A \otimes B$  is the  $n \times n'$  *block* matrix

$$A \otimes B = [A_{i,j}B],$$

with  $\ell \times \ell'$  matrix entries. We have reserved the tensor product notation for the tensor product of matrices and have eschewed the strong temptation of using  $A \otimes x_\ell$  in place of  $Ax_\ell$  when  $x_\ell$  is one of the variables.

**Proposition 6.** Suppose  $p \in \mathbb{R} \langle x \rangle$ . In increasing levels of generality,

- (1) if  $p(X) = 0$  for all  $n$  and all  $X \in (\mathbb{S}^{n \times n})^g$ , then  $p = 0$ ;
- (2) if there is a nonempty nc basic open semialgebraic set  $\mathcal{O}$  such that  $p(X) = 0$  on  $\mathcal{O}$  (meaning for every  $n$  and  $X \in \mathcal{O}(n)$ ,  $p(X) = 0$ ), then  $p = 0$ ;
- (3) there is an  $N$ , depending only upon the degree of  $p$ , so that for any  $n \geq N$  if there is an open subset  $O \subseteq (\mathbb{S}^{n \times n})^g$  with  $p(X) = 0$  for all  $X \in O$ , then  $p = 0$ .

*Proof.* See Exercises 28, 31, and 34. ■

**Exercise 7.** Use Proposition 6 to prove the following statement:

**Proposition 8.** *Suppose  $p \in \mathbb{R}\langle x \rangle$ . Show  $p(X)$  is symmetric for every  $n$  and every  $X \in (\mathbb{S}^{n \times n})^g$  if and only if  $p^\top = p$ .*

**2.3. Noncommutative convexity revisited and nc positivity.** Now we return with a bit more detail on our main theme, convexity. A symmetric polynomial  $p$  is *matrix convex*, if for each positive integer  $n$ , each pair of  $g$ -tuples  $X = (X_1, \dots, X_g)$  and  $Y = (Y_1, \dots, Y_g)$  in  $(\mathbb{S}^{n \times n})^g$  and each  $0 \leq t \leq 1$ ,

$$tp(X) + (1-t)p(Y) - p(tX + (1-t)Y) \succeq 0,$$

where, for an  $n \times n$  matrix  $A \in \mathbb{R}^{n \times n}$ , the notation  $A \succeq 0$  means  $A$  is positive semidefinite. Synonyms for matrix convex include both *nc convex*, and simply *convex*.

**Exercise 9.** Show that the definition here of (matrix) convex is equivalent to that given in equation (5) in the informal introduction to nc polynomials.

As we have already seen in the informal introduction to nc polynomials, even in one-variable, convexity in the noncommutative setting differs from convexity in the commutative case because here  $Y$  need not commute with  $X$ . Thus, although the polynomial  $x^4$  is a convex function of one real variable, it is not matrix convex. On the other hand, to verify that  $x^2$  is a matrix convex polynomial, observe that

$$\begin{aligned} tX^2 + (1-t)Y^2 & - (tX + (1-t)Y)^2 \\ & = t(1-t)(X^2 - XY - YX + Y^2) = t(1-t)(X - Y)^2 \succeq 0. \end{aligned}$$

A polynomial  $p \in \mathbb{R}\langle x \rangle$  is *matrix positive*, synonymously *nc positive* or simply *positive* if  $p(X) \succeq 0$  for all tuples  $X = (X_1, \dots, X_g) \in (\mathbb{S}^{n \times n})^g$ . A polynomial  $p$  is a *sum of squares* if there exists  $k \in \mathbb{N}$  and polynomials  $h_1, \dots, h_k$  such that

$$p = \sum_{j=1}^k h_j^\top h_j.$$

Because, for a matrix  $A$ , the matrix  $A^\top A$  is positive semidefinite, if  $p$  is a sum of squares, then  $p$  is positive. Though we will not discuss its proof in this chapter, we mention that, in contrast with the commutative case, the converse is true [Hel02, McC01].

**Theorem 10.** *If  $p \in \mathbb{R}\langle x \rangle$  is positive, then  $p$  is a sum of squares.*



As for convexity, note that  $p(x)$  is convex if and only if the polynomial  $q(x, y)$  in  $2g$  nc variables given by

$$q(x, y) = \frac{1}{2}(p(x) + p(y)) - p\left(\frac{x + y}{2}\right)$$

is positive.

**2.4. Directional derivatives vs. nc convexity and positivity.** Matrix convexity can be formulated in terms of positivity of the Hessian, just as in the case of a real variable. Thus we take a few moments to develop a very useful nc calculus.

Given a polynomial  $p \in \mathbb{R}\langle x \rangle$ , the  $\ell^{\text{th}}$  *directional derivative* of  $p$  in the “direction”  $h$  is

$$p^{(\ell)}(x)[h] := \left. \frac{d^\ell p(x + th)}{dt^\ell} \right|_{t=0}.$$

Thus  $p^{(\ell)}(x)[h]$  is the polynomial that evaluates to

$$\left. \frac{d^\ell p(X + tH)}{dt^\ell} \right|_{t=0} \quad \text{for every choice of } X, H \in (\mathbb{S}^{n \times n})^g.$$

We let  $p'(x)[h]$  denote the first derivative and the *Hessian*, denoted  $p''(x)[h]$  of  $p(x)$ , is the second directional derivative of  $p$  in the direction  $h$ .

Equivalently, the Hessian of  $p(x)$  can also be defined as the part of the polynomial

$$r(x)[h] := 2(p(x + h) - p(x))$$

in

$$\mathbb{R}\langle x \rangle[h] := \mathbb{R}\langle x_1, \dots, x_g, h_1, \dots, h_g \rangle$$

that is homogeneous of degree two in  $h$ .

If  $p'' \neq 0$ , that is, if  $p = p(x)$  is an nc polynomial of degree two or more, then the polynomial  $p''(x)[h]$  in the  $2g$  variables  $x_1, \dots, x_g, h_1, \dots, h_g$  is homogeneous of degree two in  $h$  and has degree equal to the degree of  $p$ .

**Example 11.**

(1) The Hessian of the polynomial  $p = x_1^2 x_2$  is

$$p''(x)[h] = 2(h_1^2 x_2 + h_1 x_1 h_2 + x_1 h_1 h_2).$$

(2) The Hessian of the polynomial  $f(x) = x^4$  (just one variable) is

$$f''(x)[h] = 2(h^2 x^2 + h x h x + h x^2 h + x h x h + x h^2 x + x^2 h^2).$$

NC convexity is neatly described in terms of the Hessian.

**Lemma 12.**  $p \in \mathbb{R}\langle x \rangle$  is nc convex if and only if  $p''(x)[h]$  is nc positive.

*Proof.* See Exercise 26. ■

**2.5. Symmetric, free, mixed, and classes of variables.** To this point, our variables  $x$  have been *symmetric* in the sense that, under the involution,  $x_j^\top = x_j$ . The corresponding polynomials, elements of  $\mathbb{R}\langle x \rangle$  are then the nc analog of polynomials in real variables, with evaluations at tuples in  $\mathbb{S}^{n \times n}$ . In various applications and settings it is natural to consider nc polynomials in other types of variables.

**2.5.1. Free variables.** The nc analog of polynomials in complex variables is obtained by allowing evaluations on tuples  $X$  of not necessarily symmetric matrices. In this case, the involution must be interpreted differently and the variables are called *free*.

In this setting, given the nc variables  $x = (x_1, \dots, x_g)$ , let  $x^\top = (x_1^\top, \dots, x_g^\top)$  denote another collection of nc variables. On the ring  $\mathbb{R}\langle x, x^\top \rangle$  define the involution  $^\top$  by the requiring  $x_j \mapsto x_j^\top$ ;  $x_j^\top \mapsto x_j$ ;  $^\top$  reverses the order of words; and linearity. For instance, for

$$q(x) = 1 + x_1^\top x_2 - x_2^\top x_1 \in \mathbb{R}\langle x, x^\top \rangle,$$

we have

$$q^\top(x) = 1 + x_2^\top x_1 - x_1^\top x_2.$$

Elements of  $\mathbb{R}\langle x, x^\top \rangle$  are polynomials in *free variables* and in this setting the variables themselves are *free*.

A polynomial  $p \in \mathbb{R}\langle x, x^\top \rangle$  is *symmetric* provided  $p^\top = p$ . In particular,  $q$  above is not symmetric, but

$$p = 1 + x_1^\top x_2 + x_2^\top x_1 \tag{7}$$

is.

A polynomial  $p \in \mathbb{R}\langle x, x^\top \rangle$  is *analytic* if there are no transposes; i.e., if  $p$  is a polynomial in  $x$  alone.

Elements of  $\mathbb{R}\langle x, x^\top \rangle$  are naturally evaluated on tuples  $X = (X_1, \dots, X_g) \in (\mathbb{R}^{\ell \times \ell})^g$ . For instance, if  $p$  is the polynomial in equation (7) and  $X = (X_1, X_2) \in (\mathbb{R}^{2 \times 2})^2$  where

$$X_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = X_2$$

then

$$p(X) = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

The space  $\mathbb{R}^{d \times d'} \langle x, x^\top \rangle$  is defined by analogy with  $\mathbb{R}^{d \times d'} \langle x \rangle$  and evaluation of elements in  $\mathbb{R}^{d \times d'} \langle x, x^\top \rangle$  at a tuple  $X \in (\mathbb{R}^{\ell \times \ell})^g$  is defined in the obvious way.

**Exercise 13.** State and prove analogs of Propositions 6 and 8 for  $\mathbb{R}\langle x, x^\top \rangle$  and evaluations from  $(\mathbb{R}^{\ell \times \ell})^g$ .

2.5.2. *Mixed variables.* At times it is desirable to mix free and symmetric variables. We won't introduce notation for this situation as it will generally be understood from the context. Here are some examples:

**Example 14.**

$$\begin{aligned} p(x) &= x_1^\top x_1 + x_2 + \frac{3}{4}x_1x_2x_1^\top, & x_2 &= x_2^\top; \\ \text{ric}(a_1, a_2, x) &= a_1x + xa_1^\top - xa_2a_2^\top x, & x &= x^\top, \end{aligned} \quad (8)$$

In the first case  $x_1$  is free, but  $x_2$  is symmetric; and in the second  $a_1$  and  $a_2$  are free, but  $x$  is symmetric. Two additional remarks are in order about the second polynomial. First, it is a *Riccati polynomial* ubiquitous in control theory. Second, we have separated the variables into *two classes* of variables, the  $a$  variables and the  $x$  variable(s); thus  $p \in \mathbb{R}\langle a, x = x^\top \rangle$ . In applications, the  $a$  variables can be chosen to represent known (system parameters), while the  $x$  variables are unknown(s). Of course, it could be that some of the  $a$  variables are symmetric and some free and ditto for the  $x$  variables.

**Example 15.** Various directional derivatives of  $p$  in (8) are

$$\begin{aligned} D_{x_1}p(x)[h_1] &= h_1^\top x_1 + x_1^\top h_1 + \frac{3}{4}h_1x_2x_1^\top + \frac{3}{4}x_1x_2h_1^\top, & D_{x_2}p(x)[h_2] &= h_2 + \frac{3}{4}x_1h_2x_1^\top, \\ D_xp(x)[h] &= h_1^\top x_1 + x_1^\top h_1 + h_2 + \frac{3}{4}h_1x_2x_1^\top + \frac{3}{4}x_1x_2h_1^\top + \frac{3}{4}x_1h_2x_1^\top, \end{aligned}$$

Continuing with the variable class warfare, consider the following matrix-valued example.

**Example 16.** Let

$$L(a_1, a_2, x) = \begin{bmatrix} a_1x + xa_1^\top & a_2^\top x \\ xa_2 & 1 \end{bmatrix}.$$

We consider  $L \in \mathbb{R}^{2 \times 2}\langle a, x = x^\top \rangle$ ; i.e., the  $a$  variables are free, and the  $x$ -variables symmetric. Note that  $L$  is linear in  $x$  if we consider  $a_1, a_2$  fixed. Of course, if  $a_1, a_2$  and  $x$  are all scalars, then using Schur complements tells us there is a close relation between  $L$  in this example and the Riccati of the previous example.

**2.6. Noncommutative rational functions.** While it is possible to define nc functions [Tay73, SV06, Voi04, Voi10, Pop06, Pop10, KVV+, HKM10a, HKM10b], in this section we content ourselves with a relatively informal discussion of nc rational functions [Coh95, Coh06, HVM06, KVV09].

2.6.1. *Rational functions, a gentle introduction.* Noncommutative *rational expressions* are obtained by allowing inverses of polynomials. An example is the discrete time algebraic Riccati equation (DARE)

$$r(a, x) = a_1^\top x a_1 - (a_1^\top x a_2) a_1 (a_3 + a_2^\top x a_2)^{-1} (a_2^\top x a_1) + a_4, \quad x = x^\top.$$

It is a rational expression in the free variables  $a$  and the symmetric variable  $x$ , as is  $r^{-1}$ . An example, in free variables, which arises in operator theory is

$$s(x) = x^\top (1 - x x^\top)^{-1}. \quad (9)$$

Thus, we define (scalar) *nc rational expressions* for free nc variables  $x$  by starting with nc polynomials and then applying successive arithmetic operations - addition, multiplication, and inversion. We emphasize that an expression includes the order in which it is composed and no two distinct expressions are identified, e.g.,  $(x_1) + (-x_1)$ ,  $(-1) + (((x_1)^{-1})(x_1))$ , and  $0$  are different nc rational expressions.

Evaluation on polynomials naturally extends to rational expressions. If  $r$  is a rational expression in free variables and  $X \in (\mathbb{R}^{\ell \times \ell})^g$ , then  $r(X)$  is defined - in the obvious way - as long as any inverses appearing actually exist. Indeed, our main interest is in the evaluation of a rational expression. For instance, for the polynomial  $s$  above in one free variable,  $s(X)$  is defined as long as  $I - X X^\top$  is invertible and in this case,

$$s(X) = X^\top (I - X X^\top)^{-1}.$$

Generally, a nc rational expression  $r$  can be evaluated on a  $g$ -tuple  $X$  of  $n \times n$  matrices in its *domain of regularity*,  $\text{dom } r$ , which is defined as the set of all  $g$ -tuples of square matrices of all sizes such that all the inverses involved in the calculation of  $r(X)$  exist. For example, if  $r = (x_1 x_2 - x_2 x_1)^{-1}$  then  $\text{dom } r = \{X = (X_1, X_2) : \det(X_1 X_2 - X_2 X_1) \neq 0\}$ . We assume that  $\text{dom } r \neq \emptyset$ . In other words, when forming nc rational expressions we never invert an expression that is nowhere invertible.

Two rational expressions  $r_1$  and  $r_2$  are *equivalent* if  $r_1(X) = r_2(X)$  at any  $X$  where both are defined. For instance, for the rational expression  $t$  in one free variable,

$$t(x) = (1 - x^\top x)^{-1} x^\top,$$

and  $s$  from equation (9), it is an exercise to check that  $s(X)$  is defined if and only if  $t(X)$  is and moreover in this case  $s(X) = t(X)$ . Thus  $s$  and  $t$  are equivalent rational expressions. We call an equivalence class of rational expressions a *rational function*. The set of all rational functions will be denoted by  $\mathbb{R}\langle\!\langle x \rangle\!\rangle$ .

Here is an interesting example of an nc rational function with nested inverses. It is taken from [Ber76, Theorem 6.3].

**Example 17.** Consider two free variables  $x, y$ . For any  $r \in \mathbb{R}\langle x, y \rangle$  let

$$W(r) := c(x, c(x, r)^2) \cdot c(x, c(x, r)^{-1})^{-1} \in \mathbb{R}\langle x, y \rangle. \quad (10)$$

Recall that  $c$  denotes the commutator (4). Bergman's nc rational function is given by:

$$b := W(y) \cdot W(c(x, y)) \cdot W(c(x, c(x, y))^{-1}) \cdot W(c(x, c(x, c(x, y))))^{-1} \in \mathbb{R}\langle x, y \rangle. \quad (11)$$

**Exercise 18.** Consider the function  $W$  from (10). Let  $R, X$  be  $n \times n$  matrices and assume  $c(X, c(X, R)^{-1})$  exists and is invertible. Prove:

- (1) If  $n = 2$ , then  $W(R) = 0$ .
- (2) If  $n = 3$ , then  $W(R) = \det(c(X, R))$ .

**Exercise 19.** Consider Bergman's rational function (11).

- (1) Show that on a dense set of  $2 \times 2$  matrices  $(X, Y)$ ,  $b(X, Y) = 0$ .
- (2) Prove that on a dense set of  $3 \times 3$  matrices  $(X, Y)$ ,  $b(X, Y) = 1$ .

The moral of Exercise 19 is that, unlike in the case of polynomial identities, a nc rational function that vanishes on (a dense set of)  $3 \times 3$  matrices need not vanish on (a dense set of)  $2 \times 2$  matrices.

**2.6.2. Matrices of Rational Functions;  $LDL^\top$ .** One of the main ways nc rational functions occur in systems engineering is in the manipulation of matrices of polynomials. Extremely important is the  $LDL^\top$  decomposition. Consider the  $2 \times 2$  matrix with nc entries

$$M = \begin{bmatrix} a & b^\top \\ b & c \end{bmatrix}$$

where  $a = a^\top$ . The entries themselves could be nc polynomials, or even rational functions. If  $a$  is not zero, then  $M$  has the following decomposition

$$M = LDL^\top = \begin{bmatrix} I & 0 \\ ba^{-1} & I \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & c - ba^{-1}b^\top \end{bmatrix} \begin{bmatrix} I & a^{-1}b^\top \\ 0 & I \end{bmatrix}.$$

Note that this formula holds in the case that  $c$  is itself a (square) matrix nc rational function and  $b$  (and thus  $b^\top$ ) are vector-valued nc rational functions. On the other hand, if both  $a = c = 0$ , then  $M$  is the block matrix,

$$M = \begin{bmatrix} 0 & b \\ b^\top & 0 \end{bmatrix}.$$

If  $M$  is a  $k \times k$  matrix then iterating this procedure produces a decomposition of a permutation  $\Pi M \Pi^\top$  of  $M$  of the form  $\Pi M \Pi^\top = LDL^\top$  where  $D$  and  $L$  have the form

$$D = \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & d_k & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & D_{k+1} & \cdots & 0 & 0 \\ \vdots & \cdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 0 & \cdots & D_\ell & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & E \end{bmatrix} \quad (12)$$

and  $L$  has the form,

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & \ddots & 0 & 0 & 0 & 0 & \\ * & * & 1 & 0 & 0 & 0 & 0 \\ * & * & * & I_2 & 0 & 0 & 0 \\ * & * & * & * & \ddots & 0 & 0 \\ * & * & * & * & * & I_2 & 0 \\ * & * & * & * & * & * & I_a \end{bmatrix}, \quad (13)$$

where  $d_j$  are symmetric rational functions, and the  $D_j$  are nonzero  $2 \times 2$  matrices of the form

$$D_j = \begin{bmatrix} 0 & b_j \\ b_j^\top & 0 \end{bmatrix},$$

$E$  is a square 0 matrix (possibly of size  $0 \times 0$  - so absent), and  $I_2$  is the  $2 \times 2$  identity and the  $*$ 's represent possibly nonzero rational expressions (in some cases matrices of rational functions), some of the 0s are zero matrices (of the appropriate sizes), and  $a$  is the dimension of the space that  $E$  acts upon. The permutation  $\Pi$  is necessary in cases where the procedure hits a 0 on the diagonal, necessitating a permutation to bring a nonzero diagonal entry into the “pivot” position.

**Theorem 20.** *Suppose  $M(x) \in \mathbb{R}\langle x \rangle^{\ell \times \ell}$  is symmetric, and  $\Pi M \Pi^\top = LDL^\top$  where  $L, D$  are  $\ell \times \ell$  matrices with nc rational entries as in equations (13) and (12) and  $L$  respectively. If  $n$  is a positive integer and  $X \in (\mathbb{S}^{n \times n})^g$  is in the domains of both  $L$  and  $D$ , then  $M(X)$  is positive semidefinite if and only if  $D(X)$  is positive semidefinite.*

*Proof.* The proof is an easy exercise based on the fact that a square block lower triangular matrix whose diagonal blocks are invertible is itself invertible. In this case,  $L(X)$  is block lower triangular with the  $n \times n$  identity  $I_n$  as each diagonal entry. Thus  $M(X)$  and  $D(X)$  are congruent, so have the same number of negative eigenvalues. ■

**Remark 21.** Note that if  $D$  has any  $2 \times 2$  blocks  $D_j$ , then  $D(X) \succeq 0$  if and only if each  $D_j(X) = 0$ . Thus, if  $D$  has any  $2 \times 2$  blocks, generically  $D(X)$ , and hence  $M(X)$ , is not positive semidefinite (recall we assume, without loss of generality that  $D_j$  are not zero).

**2.6.3. More on rational functions.** The matrix positivity and convexity properties of nc rational functions go just like those for polynomials. One only tests a rational function  $r$  on matrices  $X$  in its domain of regularity. The definition of directional derivatives goes as before and it is easy to compute them formally. There are issues of equivalences which we avoid here, instead referring the reader to [Coh95, KVV09] or our treatment in [HMOV06].

We emphasize that proving the assertions above takes considerable effort, because of dealing with the equivalence relation. In practice one works with rational expressions, and calculations with nc rational expressions themselves are straightforward. For instance, computing the derivative of a symmetric nc rational function  $r$  leads to an expression of the form

$$Dr(x)[h] = \text{symmetrize} \left[ \sum_{\ell=1}^k a_{\ell}(x) h b_{\ell}(x) \right],$$

where  $a_{\ell}, b_{\ell}$  are nc rational functions of  $x$ , and the symmetrization of a (not necessarily symmetric) rational expression  $s$  is  $\frac{s+s^T}{2}$ .

**2.7. Exercises.** Section 3 gives a very brief introduction on nc computer algebra and some might enjoy playing with computer algebra in working some of these exercises.

Define for use in later exercises the nc polynomials

$$\begin{aligned} p &= x_1^2 x_2^2 - x_1 x_2 x_1 x_2 - x_2 x_1 x_2 x_1 - x_2^2 x_1^2 \\ q &= x_1 x_2 x_3 + x_2 x_3 x_1 + x_3 x_1 x_2 - x_1 x_3 x_2 - x_2 x_1 x_3 - x_3 x_2 x_1 \\ s &= x_1 x_3 x_2 - x_2 x_3 x_1. \end{aligned}$$

**Exercise 22.**

- (a) What is the derivative with respect to  $x_1$  in direction  $h_1$  of  $q$  and  $s$ ?
- (b) Concerning the formal derivative with respect to  $x_1$  in direction  $h_1$ .
  - (i) Show the derivative of  $r(x_1) = x_1^{-1}$  is  $-x_1^{-1} h_1 x_1^{-1}$ .
  - (ii) What is the derivative of  $u(x_1, x_2) = x_2(1 + 2x_1)^{-1}$ ?

**Exercise 23.** Consider the polynomials  $p, q, s$  and rational functions  $r, u$  from above.

- (a) Evaluate the polynomials  $p, q, s$  on some matrices of size  $1 \times 1$ ,  $2 \times 2$  and  $3 \times 3$ .

(b) Redo part (a) for the rational functions  $r, u$ .

Try to use Mathematica or MATLAB.

**Exercise 24.** Show  $c = x_1x_2 - x_2x_1$  is not symmetric, by finding  $n$  and  $X = (X_1, X_2)$  such that  $c(X)$  is not a symmetric matrix.

**Exercise 25.** Consider the following polynomials in two and three variables, respectively:

$$\begin{aligned} h_1 &= c^2 = (x_1x_2)^2 - x_1x_2^2x_1 - x_2x_1^2x_2 + (x_2x_1)^2, \\ h_2 &= h_1x_3 - x_3h_1. \end{aligned}$$

- (a) Compute  $h_1(X_1, X_2)$  and  $h_2(X_1, X_2, X_3)$  for several choices of  $2 \times 2$  matrices  $X_j$ . What do you find? Can you formulate and prove a statement?
- (b) What happens if you plug in  $3 \times 3$  matrices into  $h_1$  and  $h_2$ ?

**Exercise 26.** Prove that a symmetric nc polynomial  $p$  is matrix convex if and only if the Hessian  $p''(x)[h]$  is matrix positive, by completing the following exercise.

Fix  $n$ , suppose  $\ell$  is a positive linear functional on  $\mathbb{S}^{n \times n}$ , and consider

$$f = \ell \circ p : (\mathbb{S}^{n \times n})^g \rightarrow \mathbb{R}.$$

- (a) Show  $f$  is convex if and only if  $\frac{d^2 f(X+tH)}{dt^2} \geq 0$  at  $t = 0$  for all  $X, H \in (\mathbb{S}^{n \times n})^g$ .

Given  $v \in \mathbb{R}^n$ , consider the linear functional  $\ell(M) := v^\top M v$  and let  $f_v = \ell \circ p$ .

- (b) *Geometric:* Fix  $n$ . Show, each  $f_v$  satisfies the convexity inequality if and only if  $p$  satisfies the convexity inequality on  $(\mathbb{S}^{n \times n})^g$ ; and
- (b) *Analytic:* show, for each  $v \in \mathbb{R}^n$ ,  $f_v''(X)[H] \geq 0$  for every  $X, H \in (\mathbb{S}^{n \times n})^g$  if and only if  $p''(X)[H] \succeq 0$  for every  $X, H \in (\mathbb{S}^{n \times n})^g$ .

**Exercise 27.** For  $n \in \mathbb{N}$  let

$$s_n = \sum_{\tau \in \text{Sym}_n} \text{sign}(\tau) x_{\tau(1)} \cdots x_{\tau(n)}$$

be a polynomial of degree  $n$  in  $n$  variables. Here  $\text{Sym}_n$  denotes the symmetric group on  $n$  elements.

- (a) Prove that  $s_4$  is a polynomial identity for  $2 \times 2$  matrices. That is, for any choice of  $2 \times 2$  matrices  $X_1, \dots, X_4$ , we have

$$s_4(X_1, \dots, X_4) = 0.$$

- (b) Fix  $d \in \mathbb{N}$ . Prove that there exists a nonzero polynomial  $p$  vanishing on all tuples of  $d \times d$  matrices.



Several of the next exercises use a version of the shift operators on Fock space. With  $g$  fixed, the corresponding *Fock space*,  $\mathcal{F} = \mathcal{F}_g$ , is the Hilbert space obtained from  $\mathbb{R}\langle x \rangle$  by declaring the words to be an orthonormal basis; i.e., if  $v, w$  are words, then

$$\langle v, w \rangle = \delta_{v,w},$$

where  $\delta_{v,w} = 1$  if  $v = w$  and is 0 otherwise. Thus  $\mathcal{F}_g$  is the closure of  $\mathbb{R}\langle x \rangle$  in this inner product. For each  $j$ , the operator  $S_j$  on  $\mathcal{F}_g$  densely defined by  $S_j p = x_j p$ , for  $p \in \mathbb{R}\langle x \rangle$  is an isometry (preserves the inner product) and hence extends to an isometry on all of  $\mathcal{F}_g$ . Of course,  $S_j$  acts on an infinite dimensional Hilbert space and thus is not a matrix.

**Exercise 28.** Given a natural number  $k$ , note that  $\mathbb{R}\langle x \rangle_k$  is a finite dimensional (and hence closed) subspace of  $\mathcal{F} = \mathcal{F}_g$ . The dimension of  $\mathbb{R}\langle x \rangle_k$  is

$$\sigma(k) = \sum_{j=0}^k g^j. \quad (14)$$

Let  $V : \mathbb{R}\langle x \rangle_k \rightarrow \mathcal{F}$  denote the inclusion and

$$T_j = V^\top S_j V.$$

Thus  $T_j$  does act on a finite dimensional space, and  $T = (T_1, \dots, T_g) \in (\mathbb{R}^{n \times n})^g$ , for  $n = \sigma(k)$ .

(a) Show, if  $v$  is a word of length at most  $k-1$ , then

$$T_j v = x_j v;$$

and  $T_j v = 0$  if the length of  $v$  is  $k$ .

(b) Determine  $T_j^\top$ ;

(c) Show, if  $p$  is a nonzero polynomial of degree at most  $k$  and  $Y_j = T_j + T_j^\top$ , then  $p(Y)\emptyset \neq 0$ ;

(d) Conclude, if, for every  $n$  and  $X \in (\mathbb{S}^{n \times n})^g$ ,  $p(X) = 0$ , then  $p$  is 0.

Exercise 28 shows there are no nc polynomials vanishing on all tuples of (symmetric) matrices of all sizes. The next exercise will lead the reader through an alternative proof inspired by standard methods of polynomial identities.

**Exercise 29.** Let  $p \in \mathbb{R}\langle x \rangle_n$  be an analytic polynomial that vanishes on  $(\mathbb{R}^{n \times n})^g$  (same fixed  $n$ ). Write  $p = p_0 + p_1 + \dots + p_n$ , where  $p_j$  is the homogeneous part of  $p$  of degree  $j$ .

(a) Show that  $p_j$  also vanishes on  $(\mathbb{R}^{n \times n})^g$ .

- (b) A polynomial  $q$  is called multilinear if it is homogeneous of degree one with respect to all of its variables. Equivalently, each of its monomials contains all variables exactly once, i.e.,

$$q = \sum_{\pi \in S_n} \alpha_\pi X_{\pi(1)} \cdots X_{\pi(n)}.$$

Using the staircase matrices  $E_{11}, E_{12}, E_{22}, E_{23}, \dots, E_{n-1,n}, E_{nn}$  show that a nonzero multilinear polynomial  $q$  of degree  $n$  cannot vanish on all  $n \times n$  matrices.

- (c) By (a) we may assume  $p$  is homogeneous. By induction on the biggest degree a variable in  $p$  can have, prove that  $p = 0$ . *Hint:* What are the degrees of the variables appearing in

$$p(x_1 + \hat{x}_1, x_2, \dots, x_g) - p(x_1, x_2, \dots, x_g) - p(\hat{x}_1, x_2, \dots, x_g)?$$

**Exercise 30.** Redo Exercise 29 for a polynomial

- (a)  $p \in \mathbb{R}\langle x, x^\top \rangle$ , not necessarily analytic, vanishing on all tuples of matrices;  
 (b)  $p \in \mathbb{R}\langle x \rangle$  vanishing on all tuples of symmetric matrices.

**Exercise 31.** Show, if  $p \in \mathbb{R}\langle x \rangle$  vanishes on a nonempty basic open semialgebraic set, then  $p = 0$ .

**Exercise 32.** Suppose  $p \in \mathbb{R}\langle x \rangle$ ,  $n$  is a positive integer and  $O \subseteq (\mathbb{S}^{n \times n})^g$  is an open set. Show, if  $p(X) = 0$  for each  $X \in O$ , then  $P(X) = 0$  for each  $X \in (\mathbb{S}^{n \times n})^g$ . *Hint:* given  $X_0 \in O$  and  $X \in (\mathbb{S}^{n \times n})^g$ , consider the matrix valued polynomial,

$$q(t) = p(X_0 + tX).$$

**Exercise 33.** Suppose  $r \in \mathbb{R}\langle x \rangle$  is a rational function and there is a nonempty basic open semialgebraic set  $\mathcal{O} \subseteq \text{dom}(r)$  with  $r|_{\mathcal{O}} = 0$ . Show that  $r = 0$ .

**Exercise 34.** Prove item (3) of Proposition 6. You may wish to use Exercises 32 and 28.

**Exercise 35.** Prove the following proposition:

**Proposition 36.** *If  $\pi : \mathbb{R}\langle x \rangle \rightarrow \mathbb{R}^{n \times n}$  is an involution preserving homomorphism, then there is an  $X \in (\mathbb{S}^{n \times n})^g$  such that  $\pi(p) = p(X)$ ; i.e., all finite dimensional representations of  $\mathbb{R}\langle x \rangle$  are evaluations.*

**Exercise 37.** Do the algebra to show

$$x^\top(1 - xx^\top)^{-1} = (1 - x^\top x)^{-1}x^\top.$$

(This is a key fact used in the model theory for contractions [NFBK10].)

**Exercise 38.** Give an example of symmetric  $2 \times 2$  matrices  $X, Y$  such that  $X \succeq Y \succeq 0$ , but  $X^2 \not\succeq Y^2$ .

This failure of a basic order property of  $\mathbb{R}$  for  $\mathbb{S}^{n \times n}$  is closely related to the rigid nature of positivity and convexity in the nc setting.

**Exercise 39.** Antiderivatives.

- (a) Is  $q(x)[h] = xh + hx$  the derivative of any nc polynomial  $p$ ? If so what is  $p$ ?
- (b) Is  $q(x)[h] = hhx + hxx + xhh$  the second derivative of any nc polynomial  $p$ ? If so what is  $p$ ?
- (c) Describe in general which polynomials  $q(x)[h]$  are the derivative of some nc polynomial  $p(x)$ .
- (d) Check you answer against the theory in [GHV11].

**Exercise 40.** (Requires background in algebra) Show that  $\mathbb{R}\langle x \rangle$  is a division ring; i.e., the nc rational functions form a ring in which every nonzero element is invertible.

**Exercise 41.** In this exercise we will establish that it is possible to embed the free algebra  $\mathbb{R}\langle x_1, \dots, x_g \rangle$  into  $\mathbb{R}\langle x, y \rangle$  for any  $g \in \mathbb{N}$ .

- (a) Show that the subalgebra of  $\mathbb{R}\langle x, y \rangle$  generated by  $xy^n$ ,  $n \in \mathbb{N}_0$ , is free.
- (b) Ditto for the subalgebra generated by

$$x_1 = x, \quad x_2 = c(x_1, y), \quad x_3 = c(x_2, y), \quad \dots, \quad x_n = c(x_{n-1}, y), \dots$$

Here, as before,  $c$  is the commutator,  $c(a, b) = ab - ba$ .

A comprehensive study of free algebras and nc rational functions from an algebraic viewpoint is developed in [Coh95, Coh06].

**Exercise 42.** As a hard exercise, numerically verify that the set

$$\text{ncTV}(2) = \{X \in (\mathbb{S}^{2 \times 2})^2 : 1 - X_1^4 - X_2^4 \succ 0\}$$

is not convex. That is, find  $X = (X_1, X_2)$  and  $Y = (Y_1, Y_2)$  where  $X_1, X_2, Y_1, Y_2$  are  $2 \times 2$  symmetric matrices such that both

$$1 - X_1^4 - X_2^4 \succ 0 \quad \text{and} \quad 1 - Y_1^4 - Y_2^4 \succ 0,$$

but

$$1 - \left( \frac{X_1 + Y_1}{2} \right)^4 - \left( \frac{X_2 + Y_2}{2} \right)^4 \not\succ 0.$$

You may wish to write a numerical search routine.

### 3. COMPUTER ALGEBRA SUPPORT

There are several computer algebra packages available to ease the first contact with free convexity and positivity. In this section we briefly describe two of them:

- (1) [NCAIgebra](#) running under Mathematica;
- (2) [NCSOSTools](#) running under MATLAB.

The former is more universal in that it implements manipulation with noncommutative variables, including nc rationals, and several algorithms pertaining to convexity. The latter is focused on nc positivity and numerics.

**3.1. NCAIgebra.** NCAIgebra [[HOMS+](#)] runs under Mathematica and gives it the capability of manipulating noncommuting algebraic expressions. An important part of the package (which we shall not go into here) is NCGB, which computes noncommutative Groebner Bases and has extensive sorting and display features as well as algorithms for automatically discarding “redundant” polynomials.

We recommend the user to have a look at the Mathematica notebook [NCBasicCommandsDemo](#) available from the NCAIgebra website

<http://math.ucsd.edu/~ncalg/>

for the basic commands and their usage in NCAIgebra. Here is a sample.

The basic ingredients are (symbolic) variables, which can be either noncommutative or commutative. At present, single-letter lower case variables are noncommutative by default and all others are commutative by default. To change this one can employ

*NCAIgebra Command:* `SetNonCommutative[listOfVariables]` to make all the variables appearing in `listOfVariables` noncommutative. The converse is given by

*NCAIgebra Command:* `SetCommutative`.

**Example 43.** Here is a sample session in Mathematica running NCAIgebra.

```
In[1]:= a ** b - b ** a
Out[1]= a ** b - b ** a
```

```
In[2]:= A ** B - B ** A
Out[2]= 0
```

```
In[3]:= A ** b - b ** a
Out[3]= A b - b ** a
```

```
In[4]:= CommuteEverything[a ** b - b ** a]
Out[4]= 0
```

```
In[5]:= SetNonCommutative[A, B]
Out[5]= {False, False}
```

```
In[6]:= A ** B - B ** A
Out[6]= A ** B - B ** A
```

```
In[7]:= SetNonCommutative[A];SetCommutative[B]
Out[7]= {True}
```

```
In[8]:= A ** B - B ** A
Out[8]= 0
```

Slightly more advanced is the `NCAgebra` command to generate the directional derivative of a polynomial  $p(x, y)$  with respect to  $x$ , which is denoted by  $D_x p(x, y)[h]$ :

*NCAgebra Command:* `DirectionalD[Function p, x, h]`, and is abbreviated

*NCAgebra Command:* `DirD`.

**Example 44.** Consider

$$a = x ** x ** y - y ** x ** y$$

Then

$$\text{DirD}[a, x, h] = (h ** x + x ** h) ** y - y ** h ** y$$

or in expanded form,

$$\text{NCExpand}[\text{DirD}[a, x, h]] = h ** x ** y + x ** h ** y - y ** h ** y$$

Note that we have used

*NCAgebra Command:* `NCExpand[Function p]` to expand a noncommutative expression. The command comes with a convenient abbreviation

*NCAgebra Command:* `NCE`.

NCAAlgebra is capable of much more. For instance, is a given noncommutative function “convex”? You type in a function of noncommutative variables; the command

*NCAAlgebra Command:* `NCCConvexityRegion[Function, ListOfVariables]` tells you where the (symbolic) Function is convex in the Variables. The algorithm comes from the paper of Camino, Helton, Skelton, Ye [CHSY03].

*NCAAlgebra Command:* `{L, D, U, P}:=NCLDUDecomposition[Matrix]`. Computes the LDU Decomposition of Matrix and returns the result as a 4 tuple. The last entry is a Permutation matrix which reveals which pivots were used. If Matrix is symmetric then  $U = L^T$ .

The NCAAlgebra website comes with extensive documentation. A more advanced notebook with a hands on demonstration of applied capabilities of the package is [DemoBRL.nb](#); it derives the Bounded Real Lemma for a linear system.

**Exercise 45.** For the polynomials and rational functions defined at the beginning of Section 2.7, use NCAAlgebra to calculate

- (a) `p**q` and `NCExpand[p**q]`
- (b) `NCCollect[p**q, x1]`
- (c) `D[p,x1,h1]` and `D[u,x1,h1]`

3.1.1. *Warning.* The Mathematica substitute commands `/.`, `/>` and `/:>` are not reliable in NCAAlgebra, so a user should use NCAAlgebra’s `Substitute` command.

**Example 46.** Here is an example of unsatisfactory behavior of the built-in Mathematica function.

```
In[1]:= (x ** a ** b) /. {a ** b -> c}
Out[1]= x ** a ** b
```

On the other hand, NCAAlgebra performs as desired:

```
In[2]:= Substitute[x ** a ** b, a ** b -> c]
Out[2]= x ** c
```

3.2. **NCSOSTools.** A reader mainly interested in positivity of noncommutative polynomials might be better served by NCSOSTools [CKP11]. NCSOSTools is an open source MATLAB toolbox for

- (a) basic symbolic computation with polynomials in noncommuting variables;

- (b) constructing and solving sum of hermitian squares (with commutators) programs for polynomials in noncommuting variables.

It is normally used in combination with standard semidefinite programming software to solve these constructed LMIs.

The NCSOSTools website

<http://ncsostools.fis.unm.si>

contains documentation and a demo notebook [NCSOSToolsdemo](#) to give the user a gentle introduction into its features.

**Example 47.** Despite some ability to manipulate symbolic expressions, MATLAB cannot handle noncommuting variables. They are implemented in NCSOSTools.

*NCSOSTools Command:* `NCvars x` introduces a noncommuting variable  $x$  into the workspace.

NCSOSTools is well equipped to work with commutators and sums of (hermitian) squares. Recall: a *commutator* is an expression of the form  $fg - gf$ .

**Exercise 48.** Use NCSOSTools to check whether the polynomial  $x^2yx + yx^3 - 2xyx^2$  is a sum of commutators. (*Hint:* Try the `NCisCycEq` command.) If so, can you find such an expression?

Let us demonstrate an example with sums of squares.

**Example 49.** Consider

$$\begin{aligned} f = & 5 + x^2 - 2x^3 + x^4 + 2xy + x*y*x*y - x*y^2 + x*y^2*x \\ & - 2*y + 2*y*x + y*x^2*y - 2*y*x*y + y*x*y*x - 3*y^2 - y^2*x + y^4 \end{aligned}$$

Is  $f$  matrix positive? By Theorem 10 it suffices to check whether  $f$  is a sum of squares. This is easily done using

*NCSOSTools Command:* `NCsos(f)`, which checks the polynomial  $f$  is a sum of squares. Running `NCsos(f)` tells us that  $f$  is indeed a sum of squares. What NCSOSTools does, is transform this question into a semidefinite program (SDP) and then calls a solver. `NCsos` comes with several options. Its full command line is

$$[\text{IsSohs}, X, \text{base}, \text{sohs}, g, \text{SDP\_data}, L] = \text{NCsos}(f, \text{params})$$

The meaning of the output is as follows:

- `IsSohs` equals 1 if the polynomial  $f$  is a sum of hermitian squares and 0 otherwise;
- `X` is the Gram matrix solution of the corresponding SDP returned by the solver;

- **base** is a list of words which appear in the SOHS decomposition;
- **sohs** is the SOHS decomposition of  $f$ ;
- **g** is the NCpoly representing  $\sum_i m_i^\top m_i$ ;
- **SDP\_data** is a structure holding all the data used in SDP solver;
- **L** is the operator representing the dual optimization problem (i.e., the dual feasible SDP matrix).

**Exercise 50.** Use NCSOSTools to compute the smallest eigenvalue  $f(X, Y)$  can attain for a pair of symmetric matrices  $(X, Y)$ . Can you also find a minimizer pair  $(X, Y)$ ?

**Exercise 51.** Let  $f = y^2 + (xy - 1)^\top(xy - 1)$ . Show that

- $f(X, Y)$  is always positive semidefinite.
- For each  $\epsilon > 0$  there is a pair of symmetric matrices  $(X, Y)$  so that the smallest eigenvalue of  $f(X, Y)$  is  $\epsilon$ .
- Can  $f(X, Y)$  be singular?

The moral of Example 51 is that even if an nc polynomial is bounded from below, it need not attain its minimum.

**Exercise 52.** Redo the Exercise 51 for  $f(x) = x^\top x + (xx^\top - 1)^\top(xx^\top - 1)$ .

#### 4. A GRAM-LIKE REPRESENTATION

The next two sections are devoted to a powerful representation of quadratic functions  $q$  in nc variables which takes a strong form when  $q$  is matrix positive; we call it a *QuadratischePositivstellensatz*. Ultimately we shall apply this to  $q(x)[h] = p''(x)[h]$  and show that if  $p$  is matrix convex (i.e.,  $q$  is matrix positive), then  $p$  has degree two. We begin by illustrating our grand scheme with examples.

##### 4.1. Illustrating the ideas.

**Example 53.** The (symmetric) polynomial  $p(x) = x_1x_2x_1 + x_2x_1x_2$  (in symmetric variables) has Hessian  $q(x)[h] = p''(x)[h]$  which is homogeneous quadratic in  $h$  and is

$$q(x)[h] = 2h_1h_2x_1 + 2h_1x_2h_1 + 2h_2h_1x_2 + 2h_2x_1h_2 + 2x_1h_2h_1 + 2x_2h_1h_2.$$

We can write  $q$  in the form

$$q(x)[h] = \begin{bmatrix} h_1 & h_2 & x_2h_1 & x_1h_2 \end{bmatrix} \begin{bmatrix} 2x_2 & 0 & 0 & 2 \\ 0 & 2x_1 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_1x_2 \\ h_2x_1 \end{bmatrix}.$$



The representation of  $q$  displayed above is of the form

$$q(x)[h] = V(x)[h]^\top Z(x)V(x)[h]$$

where  $Z$  is called the *middle matrix* (MM) and  $V$  the *border vector* (BV). The MM does not contain  $h$ . The BV is linear in  $h$  with  $h$  always on the left. In Section 4.2 we define this border vector-middle matrix (BV-MM) representation generally for nc polynomials  $q(x)[h]$  which are homogeneous of degree two in the  $h$  variables. Note the entries of the BV are distinct monomials.

**Example 54.** Let  $p = x_2x_1x_2x_1 + x_1x_2x_1x_2$ . Then

$$\begin{aligned} q = p'' &= 2h_1h_2x_1x_2 + 2h_1x_2h_1x_2 + 2h_1x_2x_1h_2 + 2h_2h_1x_2x_1 + 2h_2x_1h_2x_1 + 2h_2x_1x_2h_1 \\ &\quad + 2x_1h_2h_1x_2 + 2x_1h_2x_1h_2 + 2x_1x_2h_1h_2 + 2x_2h_1h_2x_1 + 2x_2h_1x_2h_1 + 2x_2x_1h_2h_1. \end{aligned}$$

The BV-MM representation for  $q$  is

$$q = \begin{bmatrix} h_1 & h_2 & x_2h_1 & x_1h_2 & x_1x_2h_1 & x_2x_1h_2 \end{bmatrix} \begin{bmatrix} 0 & 2x_2x_1 & 2x_2 & 0 & 0 & 2 \\ 2x_1x_2 & 0 & 0 & 2x_1 & 2 & 0 \\ 2x_1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 2x_2 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_1x_2 \\ h_2x_1 \\ h_1x_2x_1 \\ h_2x_1x_2 \end{bmatrix}$$

**Example 55.** In the one variable case with  $h_1 = h_1^\top$  we abbreviate  $h_1$  to  $h$ . Fix some nc variables not necessarily symmetric  $w := (a, b, d, e)$  and consider

$$q(w)[h] := hah + e^\top h b h + h b^\top h e + e^\top h d h e. \quad (15)$$

which is a quadratic function of  $h$ . It can be written in the BV-MM form

$$q(w)[h] = \begin{bmatrix} h & e^\top h \end{bmatrix} \begin{bmatrix} a & b^\top \\ b & d \end{bmatrix} \begin{bmatrix} h \\ h e \end{bmatrix}. \quad (16)$$

The representation is unique.

Observe (16) contrasts strongly with the commutative case wherein (15) takes the form

$$q(w)[h] = h(a + e^\top b + b^\top e + e^\top d e)h.$$

**Example 56.** The Hessian of  $p(x) = x^4$  is

$$\begin{aligned} q(x)[h] &:= p''(x)[h] = 2(x^2h^2 + xh^2x + h^2x^2) \\ &\quad + 2(xh x h + h x h x) \\ &\quad + hx^2h, \end{aligned} \quad (17)$$

a polynomial that is homogeneous of degree two in  $x$  and homogeneous of degree two in  $h$  that can be expressed as

$$q(x)[h] = 2 \begin{bmatrix} h & xh & x^2h \end{bmatrix} \begin{bmatrix} x^2 & x & 1 \\ x & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} h \\ hx \\ hx^2 \end{bmatrix}.$$

Notice that the contribution of the main antidiagonal of the MM for  $q$  in Example 56 (all 1's) corresponds to the right hand side of first line of (17). Indeed, each antidiagonal corresponds to a line of (17).

**Exercise 57.** In Example 56, for which symmetric matrices  $X$  is  $Z(X)$  positive semidefinite?

**Exercise 58.** What is the MM  $Z(x)$  for  $p(x) = x^3$ ? For which symmetric matrices  $X$  is  $Z(X)$  positive semidefinite?

**Exercise 59.** Compute middle matrix representations using [NCAlgebra](#). The command is

$$\{lt, mq, rt\} = \text{NCMatrixOfQuadratic}[q, \{h, k\}]$$

In the output  $mq$  is the MM and  $rt$  is the BV and  $lt$  is  $(rt)^\top$ . For examples, see `NCConvexityRegionDemo.nb` In the `NC/DEMOS` directory.

4.1.1. *The positivity of  $q$  vs. positivity of the MM.* In this section we let  $q(x)[h]$  denote a polynomial which is homogeneous of degree two in  $h$ , but which is not necessarily the Hessian of a nc polynomial. While we have focused on Hessians, such a  $q$  will still have a BV-MM representation. *So what good is this representation?* After all one expects that  $q$  could have wonderful properties, such as positivity, which are not shared by its middle matrix. No, the striking thing is that positivity of  $q$  implies positivity of the MM. Roughly we shall prove what we call the *Quadratische Positivstellensatz*, which is essentially Theorem 3.1 of [CHSY03](#).

**Theorem 60.** *If the polynomial<sup>2</sup>  $q(x)[h]$  is homogeneous quadratic in  $h$ , then  $q$  is matrix positive if and only if its middle matrix  $Z$  is matrix positive.*

*More generally, suppose  $\mathcal{O}$  is a nonempty nc basic open semialgebraic set. If  $q(X)[H]$  is positive semidefinite for all  $n \in \mathbb{N}$ ,  $X \in \mathcal{O}(n)$  and  $H \in (\mathbb{S}^{n \times n})^g$ , then  $Z(X) \succeq 0$  for all  $X \in \mathcal{O}$ .*

---

<sup>2</sup>This theorem is true (but not proved here) for  $q$  which are nc rational in  $x$ .

We emphasize that, in the theorem, the convention that the terms of the border vector are distinct is in force.

To foreshadow Section 5 and to give an idea of the proof of Theorem 60, we illustrate it on an example in one variable. This time we use a *free* rather than symmetric variable since proofs are a bit easier.

Consider the noncommutative quadratic function  $q$  given by

$$q(w)[h] := h^\top b h + e^\top h^\top c h + h^\top c^\top h e + e^\top h^\top a h e \quad (18)$$

where  $w = (a, b, c, e)$ . The border vector  $V(w)[h]$  and the coefficient matrix  $Z(w)$  with noncommutative entries are

$$V(w)[h] = \begin{bmatrix} h \\ h e \end{bmatrix} \quad \text{and} \quad Z(w) = \begin{bmatrix} b & c^\top \\ c & a \end{bmatrix},$$

that is,  $q$  has the form

$$q(w)[h] = V(w)[h]^\top Z(w) V(w)[h] = \begin{bmatrix} h^\top & e^\top h^\top \end{bmatrix} \begin{bmatrix} b & c^\top \\ c & a \end{bmatrix} \begin{bmatrix} h \\ h e \end{bmatrix}.$$

Now, if in equation (18) the elements  $a, b, c, e, h$  are replaced by matrices in  $\mathbb{R}^{n \times n}$ , then the noncommutative quadratic function  $q(w)[h]$  becomes a matrix valued function  $q(W)[H]$ . The matrix valued function  $q[H]$  is matrix positive if and only if  $v^\top q(W)[H] v \geq 0$  for all vectors  $v \in \mathbb{R}^n$  and all  $H \in \mathbb{R}^{n \times n}$ . Or equivalently, the following inequality must hold

$$\begin{bmatrix} v^\top H^\top & v^\top E^\top H^\top \end{bmatrix} Z \begin{bmatrix} H v \\ H E v \end{bmatrix} \geq 0. \quad (19)$$

Let

$$y^\top := \begin{bmatrix} v^\top H^\top & v^\top E^\top H^\top \end{bmatrix}.$$

Then (19) is equivalent to  $y^\top Z y \geq 0$ . Now it suffices to prove that all vectors of the form  $y$  sweep  $\mathbb{R}^{2n}$ . This will be completely analyzed in full generality in Section 5.1 but next we give the proof for our simple situation.

Suppose for a given  $v$ , with  $n \geq 2$ , the vectors  $v$  and  $Ev$  are linearly independent. Let  $y = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$  be any vector in  $\mathbb{R}^{2n}$ , then we can choose  $H \in \mathbb{R}^{n \times n}$  with the property that  $v_1 = H v$  and  $v_2 = H E v$ . It is clear that

$$\mathcal{R}^v := \left\{ \begin{bmatrix} H v \\ H E v \end{bmatrix} : H \in \mathbb{R}^{n \times n} \right\} \quad (20)$$

is all  $\mathbb{R}^{2n}$  as required.

Thus we are finished unless for all  $v$  the vectors  $v$  and  $Ev$  are linearly dependent. That is for all  $v$ ,  $\lambda_1(v)v + \lambda_2(v)Ev = 0$  for nonzero  $\lambda_1(v)$  and  $\lambda_2(v)$ . Note  $\lambda_2(v) \neq 0$ , unless  $v = 0$ . Set  $\tau(v) := \frac{\lambda_1(v)}{\lambda_2(v)}$ , then the linear dependence becomes  $\tau(v)v + Ev = 0$  for all  $v$ . It turns out that this does not happen unless  $E = \tau I$  for some  $\tau \in \mathbb{R}$ . This is a baby case of Theorem 92 which comes later and is a subject unto itself.

To finish the proof pick a  $v$  which makes  $\mathcal{R}^v$  equal all of  $\mathbb{R}^{2n}$ . Then  $v^\top q(W)[H]v \geq 0$  implies that  $Z \succeq 0$ , by (19).  $\blacksquare$

**4.2. Details of the Middle Matrix representation.** The following representation for symmetric nc polynomials  $q(x)[h]$  that are of degree  $\ell$  in  $x$  and homogeneous of degree two in  $h$  is exploited extensively in this subject:

$$q(x)[h] = \begin{bmatrix} V_0^\top & V_1^\top & \cdots & V_{\ell-1}^\top & V_\ell^\top \end{bmatrix} \begin{bmatrix} Z_{00} & Z_{01} & \cdots & Z_{0,\ell-1} & Z_{0\ell} \\ Z_{10} & Z_{11} & \cdots & Z_{1,\ell-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Z_{\ell-1,0} & Z_{\ell-2,1} & \cdots & 0 & 0 \\ Z_{\ell 0} & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{\ell-1} \\ V_\ell \end{bmatrix}, \quad (21)$$

where:

- (1) The degree  $d$  of  $q(x)[h]$  is  $d = \ell + 2$ .
- (2)  $V_j = V_j(x)[h]$ ,  $j = 0, \dots, \ell$ , is a vector of height  $g^{j+1}$  whose entries are monomials of degree  $j$  in the  $x$  variables and degree one in the  $h$  variables. The  $h$  always appears to the left. In particular,  $V(x)[h]$  is a vector of height  $g\sigma(\ell)$ , where as in (14),

$$\sigma(\ell) = 1 + g + \cdots + g^\ell.$$

- (3)  $Z_{ij} = Z_{ij}(x)$ , is a matrix of size  $g^{i+1} \times g^{j+1}$  whose entries are polynomials in the noncommuting variables  $x_1, \dots, x_g$  of degree  $\leq \ell - (i + j)$ . In particular,  $Z_{i,\ell-i} = Z_{i,\ell-i}(x)$  is a constant matrix for  $i = 0, \dots, \ell$ .
- (4)  $Z_{ij}^\top = Z_{ji}$ .

Usually the entries of the vectors  $V_j$  are ordered lexicographically.

We note that the vector of monomials,  $V(x)[h]$ , might contain monomials that are not required in the representation of the nc quadratic  $q$ . Therefore, we can omit all monomials from the border vector that are not required. This gives us a *minimal length* border vector and prevents extraneous zeros from occurring in the middle matrix. The matrix  $Z$  in the representation (21) will be referred to as the *middle matrix* (MM) of the polynomial  $q(x)[h]$  and the vectors  $V_j = V_j(x)[h]$  with monomials

as entries will be referred to as *border vectors* ( $BV$ ). It is easy to check that a minimal length border vector contains distinct monomials and once the ordering of entries of  $V$  is set the MM for a given  $q$  is unique, see Lemma 62 below.

**Example 61.** Returning to Example 54, we have for the MM representation of  $q$  that

$$V_0 = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \quad V_1 = \begin{bmatrix} h_2 x_1 \\ h_1 x_2 \end{bmatrix}, \quad V_2 = \begin{bmatrix} h_1 x_2 x_1 \\ h_2 x_1 x_2 \end{bmatrix}$$

and, for instance,

$$Z_{00} = \begin{bmatrix} 0 & 2x_2 x_1 \\ 2x_1 x_2 & 0 \end{bmatrix}, \quad Z_{01} = \begin{bmatrix} 2x^2 & 0 \\ 0 & 2x_1 \end{bmatrix}, \quad Z_{02} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}.$$

Note that generically for a polynomial  $q$  in two variables the  $V_j$  have additional terms. For instance, usually  $V_1$  is the column

$$\begin{bmatrix} h_1 x_1 \\ h_1 x_2 \\ h_2 x_1 \\ h_2 x_2 \end{bmatrix}.$$

Likewise generically  $V_2$  has eight terms. As for the  $Z_{ij}$ , for instance  $Z_{01}$  is generically  $2 \times 4$ .

**Lemma 62.** *The entries in the middle matrix  $Z(x)$  are uniquely determined by the polynomial  $q(x)[h]$  and the border vector  $V(x)[h]$ .*

*Proof.* Note every monomial in  $q(x)[h]$  has the form

$$m_L h_i m_M h_j m_R.$$

Define

$$\mathcal{R}_j := \{h_j m : m_L h_i m_M h_j m \text{ is a term in } q(x)[h]\}.$$

Given the representation  $V^\top Z V$  for  $q$ , let  $E_V$  denote the monomials in  $V$ . Then it is clear that each monomial in  $E_V$  must occur in some term of  $q$ , so it appears in  $\mathcal{R}_j$  for some  $j$ . Conversely, each term  $h_j m$  in  $\mathcal{R}_j$  corresponds to at least one term  $m_L h_i m_M h_j m$  of  $q$ , so it must be in  $E_V$ .

**Exercise 63.** Consider Equation (21) and prove the degree bound on the  $Z_{ij}$  in (3). *Hint:* Read Example 64 first.

**Example 64.** If  $p(x)$  is a symmetric polynomial of degree  $d = 4$  in  $g$  noncommuting variables, then the middle matrix  $Z(x)$  in the representation of the Hessian  $p''(x)[h]$  is

$$Z(x) = \begin{bmatrix} Z_{00}(x) & Z_{01}(x) & Z_{02}(x) \\ Z_{10}(x) & Z_{11}(x) & 0 \\ Z_{20}(x) & 0 & 0 \end{bmatrix},$$

where the block entries  $Z_{ij} = Z_{ij}(x)$  have the following structure:

- $Z_{00}$  is a  $g \times g$  matrix with nc polynomial entries of degree  $\leq 2$ ,
- $Z_{01}$  is a  $g \times g^2$  matrix with nc polynomial entries of degree  $\leq 1$ ,
- $Z_{02}$  is a  $g \times g^3$  matrix with constant entries.

All of these are proved merely by keeping track of the degrees. For example, the contribution of  $Z_{02}$  to  $p''$  is  $V_0^\top Z_{02} V_2$  whose degree is

$$\deg(V_0^\top) + \deg(Z_{02}) + \deg(V_2) = 1 + \deg(Z_{02}) + 3 \leq 4,$$

so  $\deg(Z_{02}) = 0$ .

**4.3. The Middle Matrix of  $p''$ .** The middle matrix  $Z(x)$  of the Hessian  $p''(x)[h]$  of an nc symmetric polynomial  $p(x)$  plays a key role. These middle matrices have a very rigid structure similar to that in Example 56. We illustrate with an example and then with exercises.

**Example 65.** As a warm up we first illustrate that  $Z_{02}(X) = 0$  if and only if  $Z_{11}(X) = 0$  for Example 54. To this end, observe that the contribution of the MM's extreme outer diagonal element  $Z_{02}$  to  $q$  is as follows

$$\frac{1}{2} V_0(x)[h]^\top Z_{02}(x) V_2(x)[h] = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}^\top \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} h_1 x_2 x_1 \\ h_2 x_1 x_2 \end{bmatrix} = 2h_1 h_2 x_1 x_2 + 2h_2 h_1 x_2 x_1.$$

Substitute  $h_j \rightsquigarrow x_j$  and get  $2x_1 x_2 x_1 x_2 + 2x_2 x_1 x_2 x_1$  which is  $2p(x)$ . That is,

$$p(x) = \frac{1}{2} V_0(x)[x]^\top Z_{02}(x) V_2(x)[x],$$

where  $V_k(x)[h]$  is the homogeneous, in  $x$ , of degree  $k$  part of the border vector  $V$ . Obviously,  $Z_{02} = 0$  implies  $p = 0$ .

**Exercise 66.** Show  $p(x)$  can also be obtained from  $Z_{11}$  in a similar fashion; i.e.,

$$p(x) = \frac{1}{2} V_1(x)[x]^\top Z_{11}(x) V_1(x)[x].$$

**Exercise 67.** Suppose  $p$  is homogeneous of degree  $d$  and its Hessian  $q$  has the border vector middle matrix representation  $q(x)[h] = V(x)[h]^\top Z(x) V(x)[h]$ .

(a) Show,

$$p = \frac{1}{2} V_0(x)[x]^\top Z_{0\ell} V_\ell(x)[x]$$

with  $\ell = d - 2$ . Prove this formula for  $d = 2$ ,  $d = 4$ .

(b) Show that likewise,

$$p = \frac{1}{2} V_1(x)[x]^\top Z_{1,\ell-1}(x) V_{\ell-1}(x)[x]$$

Do not cheat and look this up in [DGHM09], but do compare with Exercise 63.

**Exercise 68.** Let  $Z$  denote the middle matrix for the Hessian of a nc polynomial  $p$ . Show, if  $i + j = i' + j'$ , then  $Z_{ij} = 0$  if and only if  $Z_{i'j'} = 0$ .

**4.4. Positivity of the Middle Matrix and the demise of nc convexity.** This section focuses on positivity of the middle matrix of a Hessian.

Why should we focus on the case where  $Z(x)$  is positive semidefinite? In [HMe98] it was shown that a polynomial  $p \in \mathbb{R}\langle x \rangle$  is matrix convex if and only if its Hessian  $p''(x)[h]$  is positive (see Exercise 26). Moreover, if  $Z(x)$  is positive, then the degree of  $p(x)$  is at most two [HM04a]. The proof of this degree constraint given in Proposition 70 below using the more manageable bookkeeping scheme in this chapter, begins with the following exercise.

**Exercise 69.** Show that

$$\begin{bmatrix} A & B \\ B^\top & 0 \end{bmatrix},$$

is positive semidefinite if and only if  $A \succeq 0$  and  $B = 0$ . More refined versions of this fact appear as exercises later, see Exercise 76.

As we shall see we need not require our favorite functions be positive everywhere. It is possible to work locally, namely on an open set.

**Proposition 70.** *Let  $p = p(x)$  be a symmetric polynomial of degree  $d$  in  $g$  nc variables and let  $Z(x)$  denote the middle matrix (MM) in the BV-MM representation of the Hessian  $p''(x)[h]$ . If  $Z(X) \succeq 0$  for all  $X$  in some nonempty nc basic open semialgebraic set  $\mathcal{O}$ , then  $d$  is at most two.*

*Proof.* Arguing by contradiction, suppose  $d \geq 3$ , then  $p''(x)[h]$  is of degree  $\ell = d - 2 \geq 1$  in  $x$  and its middle matrix is of the form

$$Z = \begin{bmatrix} Z_{00} & \cdots & Z_{0\ell} \\ \vdots & \ddots & \vdots \\ Z_{\ell 0} & \cdots & 0 \end{bmatrix}.$$

Therefore,  $Z(X)$  is of the form

$$Z(X) = \begin{bmatrix} A & B \\ B^\top & 0 \end{bmatrix},$$

where  $A = A^\top$  and  $B^\top = [Z_{0\ell}(X) \ 0 \ \cdots \ 0]$ . From Exercise 67,  $p_d$ , the homogeneous degree  $d$  part of  $p$ , can be reconstructed from  $Z_{0\ell}$ . Now there is an  $X \in \mathcal{O}$  such that  $p_d(X)$  is nonzero, as otherwise  $p_d$  vanishes on a basic open semialgebraic set and is equal to 0. It follows that there is an  $X \in \mathcal{O}$  such that  $Z_{0\ell}(X)$  is not zero. Hence  $B(X)$  is not zero which implies, by Exercise 69, the contradiction that  $Z(X)$  is not positive semidefinite. ■

We have now reached our goal of showing that convex polynomials have degree  $\leq 2$ .

**Theorem 71.** *If  $p \in \mathbb{R}\langle x \rangle$  is a symmetric polynomial which is convex on a nonempty nc basic open semialgebraic set  $\mathcal{O}$ , then it has degree at most two.*

There is a version of the theorem for free variables; i.e., with  $p \in \mathbb{R}\langle x, x^\top \rangle$ .

*Proof.* The convexity of  $p$  on  $\mathcal{O}$  is equivalent to  $p''(X)[H]$  being positive semidefinite for all  $X$  in  $\mathcal{O}$ , see Exercise 26. By the QuadratischePositivstellensatz the middle matrix  $Z(x)$  for  $p''(x)[h]$  is positive on  $\mathcal{O}$ ; that is,  $Z(X) \succeq 0$  for all  $X \in \mathcal{O}$ . Proposition 70 implies degree  $p$  is at most 2. ■

**4.5. The signature of the middle matrix.** This section introduces the notion of the *signature*  $\mu_\pm(Z(x))$  of  $Z(x)$ , the middle matrix of a Hessian, or more generally a polynomial  $q(x)[h]$  which is homogeneous of degree two in  $h$ .

The *signature* of a symmetric matrix  $M$  is a triple of integers:

$$(\mu_-(M), \mu_0(M), \mu_+(M)),$$

where  $\mu_-(M)$  is the number of negative eigenvalues (counted with multiplicity);  $\mu_+(M)$  is the number of positive eigenvalues; and  $\mu_0(M)$  is the dimension of the null space of  $M$ .

**Lemma 72.** *A nc symmetric polynomial  $q(x)[h]$  homogeneous of degree two in  $h$  has middle matrix  $Z$  of the form in (21) and  $Z$  being positive semidefinite implies  $Z$  is*



of the form

$$\begin{bmatrix} Z_{00} & Z_{01} & \cdots & Z_{0, \lfloor \frac{\ell}{2} \rfloor} & 0 & \cdots \\ Z_{10} & Z_{11} & \cdots & Z_{1, \lfloor \frac{\ell}{2} \rfloor} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ Z_{\lfloor \frac{\ell}{2} \rfloor, 0} & Z_{\lfloor \frac{\ell}{2} \rfloor, 1} & \cdots & Z_{\lfloor \frac{\ell}{2} \rfloor, \lfloor \frac{\ell}{2} \rfloor} & 0 & \ddots \\ 0 & 0 & \cdots & 0 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

This lemma follows immediately from a much more general lemma.

**Lemma 73.** *If*

$$E = \begin{bmatrix} A & B & C \\ B^\top & D & 0 \\ C^\top & 0 & 0 \end{bmatrix}$$

*is a real symmetric matrix, then*

$$\mu_\pm(E) \geq \mu_\pm(D) + \text{rank } C.$$

This can be proved using the  $LDL^\top$  decomposition which we shall not do here but suggest the reader apply the  $LDL^\top$  hammer to the following simpler exercise.

#### 4.6. Exercises.

**Exercise 74.** True or False? If  $p_d$  is homogeneous of degree  $d$  and we let  $Z$  denote the middle matrix of the Hessian  $p''(x)[h]$ , then for each  $k \leq d-2$  the degree of  $Z_{i,k-i}$  is independent of  $i$ .

**Exercise 75.** Redo Exercise 26 for convexity on a nc basic open semialgebraic set.

**Exercise 76.** If  $F = \begin{bmatrix} A & C \\ C^\top & 0 \end{bmatrix}$ , then  $\mu_\pm(F) \geq \text{rank } C$ . (If you cannot do the general case, assume  $A$  is invertible.)

**Exercise 77.** If  $p(x)$  is a symmetric polynomial of degree  $d = 2$  in  $g$  noncommuting variables, then the middle matrix  $Z(x)$  in the representation of the Hessian  $p''(x)[h]$  is equal to the  $g \times g$  constant matrix  $Z_{00}$ . Substituting  $X \in (\mathbb{S}^{n \times n})^g$  for  $x$  gives

$$\mu_\pm(Z(X)) \geq \mu_\pm(Z_{00})$$

**Exercise 78.** Let  $f \in \mathbb{R}\langle x \rangle_{2d}$  and let  $V \in \langle x \rangle_d^{\sigma(d)}$  be a vector consisting of all words in  $x$  of degree  $\leq d$ . Prove:

- (a) there is a matrix  $G \in \mathbb{R}^{\sigma(d) \times \sigma(d)}$  with  $f = V^\top G V$  (any such  $G$  is called a *Gram matrix* for  $f$ );
- (b) if  $f$  is symmetric, then there is a symmetric Gram matrix for  $f$ .

**Exercise 79.** Find all Gram matrices for

- (a)  $f = x_1^4 + x_1^2 x_2 - x_1 x_2^2 + x_2 x_1^2 - x_2^2 x_1 + x_1^2 - x_2^2 + 2x_1 - x_2 + 4$ ;
- (b)  $f = c(x_1, x_2)^2$ .

**Exercise 80.** Show: if  $f \in \mathbb{R}\langle x \rangle$  is homogeneous of degree  $2d$ , then it has a unique Gram matrix  $G \in \mathbb{R}^{\sigma(d) \times \sigma(d)}$ .

**4.7. A glimpse of history.** There is a theory of operator monotone and operator convex functions which overlaps with the matrix convex functions considered here in the case of one variable. However, the points of view are substantially different, diverging markedly in several variables. Löwner introduced a class of real analytic functions in one real variable called matrix monotone functions, which we shall not define here. Löwner gave integral representations and these have developed substantially over the years. The contact with convexity came when Löwner's student Kraus [Kra36] introduced matrix convex functions  $f$  in one variable. Such a function  $f$  on  $[0, \infty) \subseteq \mathbb{R}$  can be represented as  $f(t) = tg(t)$  with  $g$  matrix monotone, so the representations for  $g$  produce representations for  $f$ . Hansen has extensive deep work on matrix convex and monotone functions whose definition in several variables is different than the one we use here, see [BT07] or [Han97]. All of this gives a beautiful integral representation characterizing matrix convex functions using techniques very different from ours. An excellent treatment of the one variable case is [Bha97, Chapter 5]. Interestingly, to the best of our knowledge, the one variable version of Theorem 71 ([HM04a]) does not seem to be explicit in this classical literature. However, it is an immediate consequence of the results of [BT07] where (not necessarily polynomial) operator convex functions on an interval are described. This and the papers of Hansen and [OST07, Uch02] are some of the more recent references in this line of convexity history orthogonal to ours.

## 5. DER QUADRATISCHEPOSITIVSTELLENSATZ

In this section we present the proof of the QuadratischePositivstellensatz (Theorem 60) which is based on the fact that local linear dependence of nc rationals (or nc polynomials) implies global linear dependence, a fact itself based on the forthcoming CHSY Lemma [CHSY03].

**5.1. The Camino, Helton, Skelton, Ye (CHSY) Lemma.** At the root of the CHSY Lemma [CHSY03] is the following linear algebra fact:

**Lemma 81.** *Fix  $n > d$ . If  $\{z_1, \dots, z_d\}$  is a linearly independent set in  $\mathbb{R}^n$ , then the codimension of*

$$\left\{ \begin{bmatrix} Hz_1 \\ Hz_2 \\ \vdots \\ Hz_d \end{bmatrix} : H \in \mathbb{S}^{n \times n} \right\} \subseteq \mathbb{R}^{nd}$$

*is  $\frac{d(d-1)}{2}$ . It is especially important that this codimension is independent of  $n$ .*

The following exercise is a variant of the Lemma 81 which is easier to prove. Thus we suggest attempting it before launching into the proof of the lemma.

**Exercise 82.** Prove if  $\{z_1, \dots, z_d\}$  is a linearly independent set in  $\mathbb{R}^n$ , then

$$\left\{ \begin{bmatrix} Hz_1 \\ Hz_2 \\ \vdots \\ Hz_d \end{bmatrix} : H \in \mathbb{R}^{n \times n} \right\} = \mathbb{R}^{nd}$$

*Hint:* it goes like the proof of (20).

*Proof of Lemma 81.* Consider the mapping  $\Phi : \mathbb{S}^{n \times n} \rightarrow \mathbb{R}^{nd}$  given by

$$H \mapsto \begin{bmatrix} Hz_1 \\ Hz_2 \\ \vdots \\ Hz_d \end{bmatrix}.$$

Since the span of  $\{z_1, \dots, z_d\}$  has dimension  $d$ , it follows that the kernel of  $\Phi$  has dimension  $\kappa = \frac{(n-d)(n-d+1)}{2}$  and hence the range has dimension  $\frac{n(n+1)}{2} - \kappa$ . To see this assertion, it suffices to assume that the span of  $\{z_1, \dots, z_d\}$  is the span of  $\{e_1, \dots, e_d\} \subseteq \mathbb{R}^n$  (the first  $d$  standard basis vectors in  $\mathbb{R}^n$ ). In this case (since  $H$  is symmetric)  $H z_j = 0$  for all  $j$  if and only if

$$H = \begin{bmatrix} 0 & 0 \\ 0 & H' \end{bmatrix},$$

where  $H'$  is a symmetric matrix of size  $(n-d) \times (n-d)$ ; in other words, this is the kernel of  $\Phi$ .

From this we deduce that the codimension of the range of  $\Phi$  is

$$nd - \left( \frac{n(n+1)}{2} - \kappa \right) = \frac{d(d-1)}{2},$$

concluding the proof. ■

Next is a straightforward extension of Lemma 81.

**Lemma 83** ([CHSY03]). *If  $n > d$  and  $\{z_1, \dots, z_d\}$  is a linearly independent subset of  $\mathbb{R}^n$ , then the codimension of*

$$\left\{ \bigoplus_{j=1}^g \begin{bmatrix} H_j z_1 \\ H_j z_2 \\ \vdots \\ H_j z_d \end{bmatrix} : H = (H_1, \dots, H_g) \in (\mathbb{S}^{n \times n})^g \right\} \subseteq \mathbb{R}^{gnd}$$

*is  $g \frac{d(d-1)}{2}$  and is independent of  $n$ .*

*Proof.* See Exercise 94. ■

Finally, the form in which we generally apply the lemma is the following.

**Lemma 84.** *Let  $v \in \mathbb{R}^n$ ,  $X \in (\mathbb{S}^{n \times n})^g$ . If the set  $\{m(X)v : m \in \langle x \rangle_d\}$  is linearly independent, then the codimension of*

$$\{V(X)[H]v : H \in (\mathbb{S}^{n \times n})^g\}$$

*is  $g \frac{\kappa(\kappa-1)}{2}$ , where  $\kappa = \sigma(d) = \sum_{j=0}^d g^j$  and where*

$$V = \bigoplus_{i=1}^g \bigoplus_{m \in \langle x \rangle_d} H_i m$$

*is the border vector associated to  $\langle x \rangle_d$ . Again, this codimension is independent of  $n$  as it only depends upon the number of variables  $g$  and the degree  $d$  of the polynomial.*

*Proof.* Let  $z_m = m(X)v$  for  $m \in \langle x \rangle_d$ . There are at most  $\kappa$  of these. Now apply the previous lemma. ■

**5.2. Linear Dependence of Symbolic Functions.** The main result in this section, Theorem 92 says roughly that if each evaluation of a set  $G_1, \dots, G_\ell$  of rational functions produces linearly dependent matrices, then they satisfy a universal linear dependence relation. We begin with a clean and easily stated consequence of Theorem 92.

In Subsection 2.1.2 we defined nc basic open semialgebraic sets. Here we define a nc basic semialgebraic set. Given matrix-valued symmetric nc polynomials  $\rho$  and  $\tilde{\rho}$ , let

$$\mathcal{D}_+^\rho(n) = \{X \in (\mathbb{S}^{n \times n})^g : \rho(X) \succ 0\},$$

and

$$\mathcal{D}^{\tilde{\rho}}(n) = \{X \in (\mathbb{S}^{n \times n})^g : \tilde{\rho}(X) \succeq 0\}.$$

Then  $\mathcal{D}$  is a nc basic semialgebraic set if there exists  $\rho_1, \dots, \rho_k$  and  $\tilde{\rho}_1, \dots, \tilde{\rho}_k$  such that  $\mathcal{D} = (\mathcal{D}(n))_{n \in \mathbb{N}}$  where

$$\mathcal{D}(n) = \left( \bigcap_j \mathcal{D}_+^{\rho_j}(n) \right) \cap \left( \bigcap_j \mathcal{D}^{\tilde{\rho}_j}(n) \right).$$

**Theorem 85.** *Suppose  $G_1, \dots, G_\ell$  are rational expressions and  $\mathcal{D}$  is a nonempty nc basic semialgebraic set on which each  $G_j$  is defined. If, for each  $X \in \mathcal{D}(n)$  and vector  $v \in \mathbb{R}^n$  the set  $\{G_j(X)v : j = 1, 2, \dots, \ell\}$  is linearly dependent, then the set  $\{G_j(X) : j = 1, 2, \dots, \ell\}$  is linearly dependent on  $\mathcal{D}$ , i.e. there exists a nonzero  $\lambda \in \mathbb{R}^\ell$  such that*

$$0 = \sum_{j=1}^{\ell} \lambda_j G_j(X) \quad \text{for all } X \in \mathcal{D}.$$

*If, in addition,  $\mathcal{D}$  contains an  $\epsilon$ -neighborhood of 0 for some  $\epsilon > 0$ , then there exists a nonzero  $\lambda \in \mathbb{R}^\ell$  such that*

$$0 = \sum_{j=1}^{\ell} \lambda_j G_j.$$

**Corollary 86.** *Suppose  $G_1, \dots, G_\ell$  are rational expressions. If, for each  $n \in \mathbb{N}$ ,  $X \in (\mathbb{S}^{n \times n})^g$ , and vector  $v \in \mathbb{R}^n$  the set  $\{G_j(X)v : j = 1, 2, \dots, \ell\}$  is linearly dependent, then the set  $\{G_j : j = 1, 2, \dots, \ell\}$  is linearly dependent, i.e., there exists a nonzero  $\lambda \in \mathbb{R}^\ell$  such that*

$$\sum_{j=1}^{\ell} \lambda_j G_j = 0.$$

**Corollary 87.** *Suppose  $G_1, \dots, G_\ell$  are rational expressions. If, for each  $n \in \mathbb{N}$  and  $X \in (\mathbb{S}^{n \times n})^g$ , the set  $\{G_j(X) : j = 1, 2, \dots, \ell\}$  is linearly dependent, then the set  $\{G_j : j = 1, 2, \dots, \ell\}$  is linearly dependent.*

The point is that the  $\lambda_j$  are independent of  $X$ . Before proving Theorem 85 we shall introduce some terminology pursuant to our more general result.

5.2.1. *Direct Sums.* We present some definitions about direct sum and sets which respect direct sums, since they are important tools.

**Definition 88.** Our definition of the *direct sum* is the usual one. Given pairs  $(X_1, v_1)$  and  $(X_2, v_2)$  where  $X_j$  are  $n_j \times n_j$  matrices and  $v_j \in \mathbb{R}^{n_j}$ ,

$$(X_1, v_1) \oplus (X_2, v_2) = (X_1 \oplus X_2, v_1 \oplus v_2)$$

where

$$X_1 \oplus X_2 := \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \quad v_1 \oplus v_2 := \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

We extend this definition to  $\mu$  terms,  $(X_1, v_1), \dots, (X_\mu, v_\mu)$  in the expected way.

In the definition below, we consider a set  $\mathcal{B}$  which is the sequence

$$\mathcal{B} := (\mathcal{B}(n)),$$

where each  $\mathcal{B}(n)$  is a set whose members are pairs  $(X, v)$  where  $X$  is in  $(\mathbb{S}^{n \times n})^g$  and  $v \in \mathbb{R}^n$ .

**Definition 89.** The set  $\mathcal{B}$  is said to *respect direct sums* if  $(X^j, v^j)$  with  $X^j \in (\mathbb{S}^{n_j \times n_j})^g$  and  $v^j \in \mathbb{R}^{n_j}$  for  $j = 1, \dots, \mu$  being contained in the set  $\mathcal{B}(n_j)$  implies that the direct sum

$$(X^1 \oplus \dots \oplus X^\mu, v^1 \oplus \dots \oplus v^\mu) = (\oplus_{j=1}^\mu X^j, \oplus_{j=1}^\mu v^j)$$

is also contained in  $\mathcal{B}(\sum n_j)$ .

**Definition 90.** By a *natural map*  $G$  on  $\mathcal{B}$ , we mean a sequence of functions  $G(n) : \mathcal{B}(n) \rightarrow \mathbb{R}^n$ , which *respects direct sums* in the sense that, if  $(X^j, v^j) \in \mathcal{B}(n_j)$  for  $j = 1, 2, \dots, \mu$ , then

$$G(\sum_1^\mu n_j)(\oplus X^j, \oplus v^j) = \oplus_1^\mu G(n_j)(X^j, v^j).$$

Typically we omit the argument  $n$ , writing  $G(X)$  instead of  $G(n)(X)$ .

Examples of sets which respect direct sums and of natural maps are provided by the following example.

**Example 91.** Let  $\rho$  be a rational expression.

- (1) The set  $\mathcal{B}^\rho = \{(X, v) : X \in \mathcal{D}^\rho \cap (\mathbb{S}^{n \times n})^g, v \in \mathbb{R}^n, n \in \mathbb{N}\}$  respects direct sums.
- (2) If  $G$  is a matrix-valued nc rational expression whose domain contains  $\mathcal{D}^\rho$ , then  $G$  determines a natural map on  $\mathcal{B}(\rho)$  by  $G(n)(X, v) = G(X)v$ . In particular, every nc polynomial determines a natural map on every nc basic semialgebraic set  $\mathcal{B}$ .

### 5.2.2. Main Result on Linear Dependence.

**Theorem 92.** *Suppose  $\mathcal{B}$  is a set which respects direct sums and  $G_1, \dots, G_\ell$  are natural maps on  $\mathcal{B}$ . If for each  $(X, v) \in \mathcal{B}$  the set  $\{G_1(X, v), \dots, G_\ell(X, v)\}$  is linearly dependent, then there exists a nonzero  $\lambda \in \mathbb{R}^\ell$  so that*

$$0 = \sum_{j=1}^{\ell} \lambda_j G_j(X, v)$$

for every  $(X, v) \in \mathcal{B}$ . We emphasize that  $\lambda$  is independent of  $(X, v)$ .

Before proving 92, we use it to prove an important earlier theorem.

*Proof of Theorem 85.* Let  $\mathcal{B}$  be given by

$$\mathcal{B}(n) = \{(X, v) : X \in \mathcal{D}^p \cap (\mathbb{S}^{n \times n})^g \text{ and } v \in \mathbb{R}^n\}.$$

Let  $G_j$  denote the natural maps,  $G_j(X, v) = G_j(X)v$ . Then  $\mathcal{B}$  and  $G_1, \dots, G_\ell$  satisfy the hypothesis of Theorem 92 and so the first conclusion of Theorem 85 follows.

The last conclusion follows because an nc rational function  $r$  vanishing on an nc basic open semialgebraic set is 0 on all  $\text{dom}(r)$  and hence is zero, cf. Exercise 33. ■

### 5.2.3. Proof of Theorem 92.

We start with a finitary version of Theorem 92:

**Lemma 93.** *Let  $\mathcal{B}$  and  $G_i$  be as in Theorem 92. If  $\mathcal{R}$  is a finite subset of  $\mathcal{B}$ , then there exists a nonzero  $\lambda(\mathcal{R}) \in \mathbb{R}^\ell$  such that*

$$\sum_{j=1}^{\ell} \lambda(\mathcal{R})_j G_j(X)v = 0,$$

for every  $(X, v) \in \mathcal{R}$ .

*Proof.* The proof relies on taking direct sums of matrices. Write the set  $\mathcal{R}$  as

$$\mathcal{R} = \{(X^1, v^1), \dots, (X^\mu, v^\mu)\},$$

where each  $(X^i, v^i) \in \mathcal{B}$ . Since  $\mathcal{B}$  respects direct sums,

$$(X, v) = (\oplus_{\nu=1}^{\mu} X^\nu, \oplus_{\nu=1}^{\mu} v^\nu) \in \mathcal{B}.$$

Hence, there exists a nonzero  $\lambda(\mathcal{R}) \in \mathbb{R}^\ell$  such that

$$0 = \sum_{j=1}^{\ell} \lambda(\mathcal{R})_j G_j(X, v).$$

Since each  $G_j$  respects direct sums, the desired conclusion follows. ■

*Proof of Theorem 92.* The proof is essentially a compactness argument, based on Lemma 93. Let  $\mathbb{B}$  denote the unit sphere in  $\mathbb{R}^\ell$ .

To  $(X, v) \in \mathcal{B}$  associate the set

$$\Omega_{(X,v)} = \left\{ \lambda \in \mathbb{B} : \lambda \cdot G(X)v = \sum_j \lambda_j G_j(X, v) = 0 \right\}.$$

Since  $(X, v) \in \mathcal{B}$ , the hypothesis on  $\mathcal{B}$  says  $\Omega_{(X,v)}$  is nonempty. It is evident that  $\Omega_{(X,v)}$  is a closed subset of  $\mathbb{B}$  and is thus compact.

Let  $\Omega := \{\Omega_{(X,v)} : (X, v) \in \mathcal{B}\}$ . Any finite sub-collection from  $\Omega$  has the form  $\{\Omega_{(X,v)} : (X, v) \in \mathcal{R}\}$  for some finite subset  $\mathcal{R}$  of  $\mathcal{B}$ , and so by Lemma 93 has a nonempty intersection. In other words,  $\Omega$  has the finite intersection property. The compactness of  $\mathbb{B}$  implies that there is a  $\lambda \in \mathbb{B}$  which is in every  $\Omega_{(X,v)}$ . This is the desired conclusion of the theorem.  $\blacksquare$

**5.3. Proof of the QuadratischePositivstellensatz.** We are now ready to give the proof of Theorem 60. Accordingly, let  $\mathcal{O}$  be a given basic open semialgebraic set. Suppose

$$q(x)[h] = V(x)[h]^\top Z(x) V(x)[h], \quad (22)$$

where  $V$  is the border vector and  $Z$  is the middle matrix; cf. (21). Clearly, if  $Z$  is matrix-positive on  $\mathcal{O}$ , then  $q(X)[H]$  is positive semidefinite for each  $n$ , each  $X \in \mathcal{O}(n)$  and  $H \in (\mathbb{S}^{n \times n})^g$ .

The converse is less trivial and requires the CHSY Lemma plus our main result on linear dependence of nc rational functions. Let  $\ell$  denote the degree of  $q(x)[h]$  in the variable  $x$ . In particular, the border vector in the representation of  $q(x)[h]$  itself has degree  $\ell$  in  $x$ . Recall  $\sigma_\ell$  from Exercise 28.

Suppose for some  $s$  and  $g$ -tuple of symmetric matrices  $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_g) \in \mathcal{O}(s)$ , the matrix  $Z(\tilde{X})$  is not positive semidefinite. By Lemma 84 and Theorem 85, there is an  $t$ , a  $Y \in \mathcal{O}(t)$ , and a vector  $\eta$  so that  $\{m(Y)\eta : m \in \langle x \rangle_\ell\}$  is linearly independent. Let  $X = \tilde{X} \oplus Y$  and  $\gamma = 0 \oplus \eta \in \mathbb{R}^{s+t}$ . Then  $Z(X)$  is not positive semidefinite and  $\{m(X)\gamma : m \in \langle x \rangle_\ell\}$  is linearly independent.

Let  $N = g \frac{\kappa(\kappa-1)}{2} + 1$ , where  $\kappa$  is given in Lemma 84 and let  $n = (s+t)N$ . Consider  $W = X \otimes I_N = (X_1 \otimes I_N, \dots, X_g \otimes I_N)$  and vector  $\omega = \gamma \otimes e$ , for any nonzero vector  $e \in \mathbb{R}^{N+1}$ . The set  $\{m(W)\omega : m \in \langle x \rangle_\ell\}$  is linearly independent and thus by Lemma 84, the codimension of  $\mathcal{M} = \{V(W)[H]\omega : H \in (\mathbb{S}^{n \times n})^g\}$  is at most  $N-1$ . On the other hand, because  $Z(X)$  has a negative eigenvalue, the matrix  $Z(W)$  has an eigenspace  $\mathcal{E}$ , corresponding to a negative eigenvalue, of dimension at



least  $N$ . It follows that  $\mathcal{E} \cap \mathcal{M}$  is nonempty; i.e., there is an  $H \in (\mathbb{S}^{n \times n})^g$  such that  $V(W)[H]\omega \in \mathcal{E}$ . In particular, this together with (22) implies

$$\langle q(W)[H]\omega, \omega \rangle = \langle Z(W)V(W)[H]\omega, V(W)\omega \rangle < 0$$

and thus,  $q(W)[H]$  is not positive semidefinite. ■

#### 5.4. Exercises.

**Exercise 94.** Prove Lemma 83.

**Exercise 95.** Let  $A \in \mathbb{R}^{n \times n}$  be given. Show, if the rank of  $A$  is  $r$ , then the matrices  $A, A^2, \dots, A^{r+1}$  are linearly dependent.

In the next exercise employ the Fock space (see Section 2.7) to prove a strengthening of Corollary 86 for nc polynomials.

**Exercise 96.** Suppose  $p_1, \dots, p_\ell \in \mathbb{R}\langle x \rangle_k$  are nc polynomials. Show, if the set of vectors

$$\{p_1(X)v, \dots, p_\ell(X)v\} \tag{23}$$

is linearly dependent for every  $(X, v) \in (\mathbb{S}^{\sigma \times \sigma})^g \times \mathbb{R}^\sigma$ , where  $\sigma = \sigma(k) = \dim \mathbb{R}\langle x \rangle_k$ , then  $\{p_1, \dots, p_\ell\}$  is linearly dependent.

**Exercise 97.** Redo Exercise 96 under the assumption that the vectors (23) are linearly dependent for all  $(X, v) \in O \times \mathbb{R}^\sigma$ , where  $O \subseteq (\mathbb{S}^{\sigma \times \sigma})^g$  is a nonempty open set.

For a more algebraic view of the linear dependence of nc polynomials we refer to [BK13].

**Exercise 98.** Prove that  $f \in \mathbb{R}\langle x \rangle$  is a sum of squares if and only if it has a positive semidefinite Gram matrix. Are then all of  $f$ 's Gram matrices positive semidefinite?

## 6. NC VARIETIES WITH POSITIVE CURVATURE HAVE DEGREE TWO

This section looks at noncommutative varieties and their geometric properties. We see a very strong rigidity when they have positive curvature which generalizes what we have already seen about convex polynomials (their graph is a positively curved variety) having degree two.

In the classical setting of a surface defined by the zero set

$$\nu(p) = \{x \in \mathbb{R}^g : p(x) = 0\}$$

of a polynomial  $p = p(x_1, \dots, x_g)$  in  $g$  commuting variables, the second fundamental form at a smooth point  $x_0$  of  $\nu(p)$  is the quadratic form,

$$h \mapsto -\langle (\text{Hess } p)(x_0)h, h \rangle, \quad (24)$$

where  $\text{Hess } p$  is the Hessian of  $p$ , and  $h \in \mathbb{R}^g$  is in the tangent space to the surface  $\nu(p)$  at  $x_0$ ; i.e.,  $\nabla p(x_0) \cdot h = 0$ .<sup>3</sup>

We shall show that in the noncommutative setting the zero set  $\mathcal{V}(p)$  of a noncommutative polynomial  $p$  (subject to appropriate irreducibility constraints) having positive curvature (even in a small neighborhood) implies that  $p$  is convex - and thus,  $p$  has degree at most two - and  $\mathcal{V}(p)$  has positive curvature everywhere; see Theorem 103 for the precise statements.

In fact there is a natural notion of the signature  $C_{\pm}(\mathcal{V}(p))$  of a variety  $\mathcal{V}(p)$  and the bound

$$\deg(p) \leq 2C_{\pm}(\mathcal{V}(p)) + 2$$

on the degree of  $p$  in terms of the signature  $C_{\pm}(\mathcal{V}(p))$  was obtained in [DHM07b]. The convention that  $C_+(\mathcal{V}(p)) = 0$  corresponds to positive curvature, since in our examples, defining functions  $p$  are typically concave or quasiconcave. One could consider characterizing  $p$  for which  $C_{\pm}(\mathcal{V}(p))$  satisfies less restrictive hypothesis than equal zero and this has been done to some extent in [DGHM09]; however, this higher level of generality is beyond our focus here. Since our goal is to present the basic ideas, we stick to positive curvature.

**6.1. NC varieties and their curvature.** We next define a number of basic geometric objects associated to the nc variety determined by an nc polynomial  $p$ .

**6.1.1. Varieties, tangent planes, and the second fundamental form.** The *variety* (zero set) of a  $p \in \mathbb{R}\langle x \rangle$  is

$$\mathcal{V}(p) := \bigcup_{n \geq 1} \mathcal{V}_n(p),$$

where

$$\mathcal{V}_n(p) := \{(X, v) \in (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n : p(X)v = 0\}.$$

---

<sup>3</sup>The choice of the minus sign in (24) is somewhat arbitrary. Classically the sign of the second fundamental form is associated with the choice of a smoothly varying vector that is normal to  $\nu(p)$ . The zero set  $\nu(p)$  has positive curvature at  $x_0$  if the second fundamental form is either positive semidefinite or negative semidefinite at  $x_0$ . For example, if we define  $\nu(p)$  using a concave function  $p$ , then the second fundamental form is negative semidefinite, while for the same set  $\nu(-p)$  the second fundamental form is positive semidefinite.

The *clamped tangent plane* to  $\mathcal{V}(p)$  at  $(X, v) \in \mathcal{V}_n(p)$  is

$$\mathcal{T}_p(X, v) := \{H \in (\mathbb{S}^{n \times n})^g : p'(X)[H]v = 0\}.$$

The *clamped second fundamental form* for  $\mathcal{V}(p)$  at  $(X, v) \in \mathcal{V}_n(p)$  is the quadratic form

$$\mathcal{T}_p(X, v) \rightarrow \mathbb{R}, \quad H \mapsto -\langle p''(X)[H]v, v \rangle.$$

Note that

$$\{X \in (\mathbb{S}^{n \times n})^g : (X, v) \in \mathcal{V}(p) \text{ for some } v \neq 0\} = \{X \in (\mathbb{S}^{n \times n})^g : \det(p(X)) = 0\}$$

is a variety in  $(\mathbb{S}^{n \times n})^g$  and typically has a *true* (commutative) tangent plane at many points  $X$ , which of course has codimension one, whereas the clamped tangent plane at a typical point  $(X, v) \in \mathcal{V}_n(p)$  has codimension on the order of  $n$  and is contained inside the true tangent plane.

6.1.2. *Full rank points.* The point  $(X, v) \in \mathcal{V}(p)$  is a *full rank point* of  $p$  if the mapping

$$(\mathbb{S}^{n \times n})^g \rightarrow \mathbb{R}^n, \quad H \mapsto p'(X)[H]v$$

is onto. The full rank condition is a nonsingularity condition which amounts to a smoothness hypothesis. Such conditions play a major role in real algebraic geometry, see [BCR98, §3.3].

As an example, consider the classical real algebraic geometry case of  $n = 1$  (and thus  $X \in \mathbb{R}^g$ ) with the commutative polynomial  $\check{p}$  (which can be taken to be the *commutative collapse* of the polynomial  $p$ ). In this case, a full rank point  $(X, 1) \in \mathbb{R}^g \times \mathbb{R}$  is a point at which the gradient of  $\check{p}$  does not vanish. Thus,  $X$  is a nonsingular point for the zero variety of  $\check{p}$ .

Some perspective for  $n > 1$  is obtained by counting dimensions. If  $(X, v) \in (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n$ , then  $H \mapsto p'(X)[H]v$  is a linear map from the  $g(n^2 + n)/2$  dimensional space  $(\mathbb{S}^{n \times n})^g$  into the  $n$  dimensional space  $\mathbb{R}^n$ . Therefore, the codimension of the kernel of this map is no bigger than  $n$ . This codimension is  $n$  if and only if  $(X, v)$  is a full rank point and in this case the clamped tangent plane has codimension  $n$ .

6.1.3. *Positive curvature.* As noted earlier, a notion of positive (really nonnegative) curvature can be defined in terms of the clamped second fundamental form.

The variety  $\mathcal{V}(p)$  has *positive curvature* at  $(X, v) \in \mathcal{V}(p)$  if the clamped second fundamental form is nonnegative at  $(X, v)$ ; i.e., if

$$-\langle p''(X)[H]v, v \rangle \geq 0 \quad \text{for every } H \in \mathcal{T}_p(X, v).$$

6.1.4. *Irreducibility: The minimum degree defining polynomial condition.* While there is no tradition of what is an effective notion of irreducibility for nc polynomials, there is a notion of minimal degree nc polynomial which is appropriate for the present context. In the commutative case the polynomial  $\check{p}$  on  $\mathbb{R}^g$  is a *minimal degree defining polynomial* for  $\nu(\check{p})$  if there does not exist a polynomial  $q$  of lower degree such that  $\nu(\check{p}) = \nu(q)$ . This is a key feature of irreducible polynomials.

**Definition 99.** A symmetric nc polynomial  $p$  is a *minimum degree defining polynomial* for a nonempty set  $\mathcal{D} \subseteq \mathcal{V}(p)$  if whenever  $q \neq 0$  is another (not necessarily symmetric) nc polynomial such that  $q(X)v = 0$  for each  $(X, v) \in \mathcal{D}$ , then

$$\deg(q) \geq \deg(p).$$

Note this contrasts with [DHM07a], where minimal degree meant a slightly weaker inequality holds.

The reader who is so inclined can simply choose  $\mathcal{D} = \mathcal{V}(p)$  or  $\mathcal{D}$  equal to the full rank points of  $\mathcal{V}(p)$ .

Now we give an example to illustrate these ideas.

6.2. **A very simple example.** In the following example, the null space

$$\mathcal{T} = \mathcal{T}_p(X, v) = \{H \in (\mathbb{S}^{n \times n})^g : p'(X)[H]v = 0\}$$

is computed for certain choices of  $p$ ,  $X$ , and  $v$ . Recall that if  $p(X)v = 0$ , then the subspace  $\mathcal{T}$  is the *clamped tangent plane* introduced in Subsection 6.1.1.

**Example 100.** Let  $X \in \mathbb{S}^{n \times n}$ ,  $v \in \mathbb{R}^n$ ,  $v \neq 0$ , let  $p(x) = x^k$  for some integer  $k \geq 1$ . Suppose that  $(X, v) \in \mathcal{V}(p)$ , that is,  $X^k v = 0$ . Then, since

$$X^k v = 0 \iff Xv = 0 \quad \text{when } X \in \mathbb{S}^{n \times n},$$

it follows that  $p$  is a minimum degree defining polynomial for  $\mathcal{V}(p)$  if and only if  $k = 1$ .

It is readily checked that

$$(X, v) \in \mathcal{V}(p) \implies p'(X)[H]v = X^{k-1}Hv,$$

and hence that  $X$  is a full rank point for  $p$  if and only if  $X$  is invertible.

Now suppose  $k \geq 2$ . Then,

$$\langle p''(X)[H]v, v \rangle = 2\langle HX^{k-2}Hv, v \rangle.$$

Therefore, if  $k > 2$

$$(X, v) \in \mathcal{V}(p) \quad \text{and} \quad p'(X)[H]v = 0 \implies XHv = 0, \text{ and so}$$

$$\langle p''(X)[H]v, v \rangle = 0.$$

To count the dimension of  $\mathcal{T}$  we can suppose without loss of generality that

$$X = \begin{bmatrix} 0 & 0 \\ 0 & Y \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top,$$

where  $Y \in \mathbb{S}^{(n-1) \times (n-1)}$  is invertible. Then, for the simple case under consideration,

$$\mathcal{T} = \{H \in \mathbb{S}^{n \times n} : h_{21}, \dots, h_{n1} = 0\},$$

where  $h_{ij}$  denotes the  $ij$  entry of  $H$ . Thus,

$$\dim \mathcal{T} = \frac{n^2 + n}{2} - (n - 1),$$

i.e.,  $\text{codim } \mathcal{T} = n - 1$ .

**Remark 101.** We remark that

$$X^k v = 0 \quad \text{and} \quad \langle p''(X)[H]v, v \rangle = 0 \implies p'(X)[H]v = 0 \quad \text{if } k = 2t \geq 4,$$

as follows easily from the formula

$$\langle p''(X)[H]v, v \rangle = 2\langle X^{t-1}Hv, X^{t-1}Hv \rangle.$$

**Exercise 102.** Let  $A \in \mathbb{S}^{n \times n}$  and let  $\mathcal{U}$  be a maximal strictly negative subspace of  $\mathbb{R}^n$  with respect to the quadratic form  $\langle Au, u \rangle$ . Prove: there exists a complementary subspace  $\mathcal{V}$  of  $\mathcal{U}$  in  $\mathbb{R}^n$  such that  $\langle Av, v \rangle \geq 0$  for every  $v \in \mathcal{V}$ .

### 6.3. Main Result: Positive curvature and the degree of $p$ .

**Theorem 103.** *Let  $p$  be a symmetric nc polynomial in  $g$  symmetric variables, let  $\mathcal{O}$  be a nc basic open semialgebraic set and let  $\mathcal{R}$  denote the full rank points of  $p$  in  $\mathcal{V}(p) \cap \mathcal{O}$ . If*

- (1)  $\mathcal{R}$  is nonempty;
- (2)  $\mathcal{V}(p)$  has positive curvature at each point of  $\mathcal{R}$ ; and
- (3)  $p$  is a minimum degree defining polynomial for  $\mathcal{R}$ ,

*then  $\deg(p)$  is at most two and  $p$  is concave.*

**6.4. Ideas and proofs.** Our aim is to give the idea behind the proof of Theorem 103 under much stronger hypotheses. We saw earlier the positivity of a quadratic on a nc basic open set  $\mathcal{O}$  imparts positivity to its MM there. The following shows this happens for thin sets (nc varieties) too. Thus, the following theorem generalizes the QuadratischePositivstellensatz, Theorem 60.

**Theorem 104.** *Let  $p, \mathcal{O}, \mathcal{R}$  be as in Theorem 103. Let  $q(x)[h]$  be a polynomial which is quadratic in  $h$  having MM representation  $q = V^\top ZV$  for which  $\deg(V) \leq \deg(p)$ . If*

$$v^\top q(X)[H]v \geq 0 \quad \text{for all } (X, v) \in \mathcal{R} \text{ and all } H, \quad (25)$$

*then  $Z(X)$  is positive semidefinite for all  $X$  with  $(X, v) \in \mathcal{R}$ .*

*Proof.* The proof of this theorem follows the proof of the QuadratischePositivstellensatz, modified to take into account the set  $\mathcal{R}$ .

Suppose for each  $(X, v) \in \mathcal{R}$  there is a linear combination  $G_{(X,v)}(x)$  of the words  $\{m(x) : \deg(m) < \deg(p)\}$  with  $G_{(X,v)}(X)v = 0$  for all  $(X, v) \in \mathcal{R}$ . Then by Theorem 92 (note that  $\mathcal{R}$  is closed under direct sums), there is a linear combination  $G \in \mathbb{R}\langle x \rangle_{\deg(p)-1}$  with  $G(X)v = 0$ . However, this is absurd by the minimality of  $p$ . Hence there is an  $(Y, v) \in \mathcal{R}$  such that  $\{m(Y)v : \deg(m) < \deg(p)\}$  is linearly independent.

Assume for some  $g$ -tuple of symmetric matrices  $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_g)$ , there is a vector  $\tilde{v}$  such that  $(\tilde{X}, \tilde{v}) \in \mathcal{R}$ , and the matrix  $Z(\tilde{X})$  is not positive semidefinite. Let  $X = \tilde{X} \oplus Y$  and  $\gamma = \tilde{v} \oplus v$ . Then  $(X, \gamma) \in \mathcal{R}(\ell)$  for some  $\ell$ ; the matrix  $Z(X)$  is not positive semidefinite; and  $\{m(X)\gamma : \deg(m) < \deg(p)\}$  is linearly independent.

Let  $N = g \frac{\kappa(\kappa-1)}{2} + 1$ , where  $\kappa$  is given in Lemma 84 and let  $n = \ell N$ . Consider  $W = X \otimes I_N = (X_1 \otimes I_N, \dots, X_g \otimes I_N)$  and vector  $\omega = \gamma \otimes e$ , where  $e \in \mathbb{R}^N$  is the vector with each entry equal to 1. Then,  $(W, \omega) \in \mathcal{R}(n)$ , and the set  $\{m(W)\omega : m \in \langle x \rangle_\ell\}$  is linearly independent and thus by Lemma 84, the codimension of  $\mathcal{M} = \{V(W)[H]\omega : H \in (\mathbb{S}^{n \times n})^g\}$  is at most  $N - 1$ . On the other hand, because  $Z(X)$  has a negative eigenvalue, the matrix  $Z(W)$  has an eigenspace  $\mathcal{E}$ , corresponding to a negative eigenvalue, of dimension at least  $N$ . It follows that  $\mathcal{E} \cap \mathcal{M}$  is nonempty; i.e., there is an  $H \in (\mathbb{S}^{n \times n})^g$  such that  $V(W)[H]\omega \in \mathcal{E}$ . In particular,

$$\langle q(W)[H]\omega, \omega \rangle = \langle Z(W)V(W)[H]\omega, V(W)\omega \rangle < 0$$

and thus,  $q(W)[H]$  is not positive semidefinite. ■

**6.4.1. The modified Hessian.** Our main tool for analyzing the curvature of noncommutative varieties is a variant of the Hessian for symmetric nc polynomials  $p$ . The curvature of  $\mathcal{V}(p)$  is defined in terms of  $\text{Hess}(p)$  compressed to tangent planes, for

each dimension  $n$ . This compression of the Hessian is awkward to work with directly, and so we associate to it a quadratic polynomial  $q(x)[h]$  carrying all of the information of  $p''$  compressed to the tangent plane, but having the key property (25). We shall call this  $q$  we construct the relaxed Hessian. The first step in constructing the relaxed Hessian is to consider the simpler *modified Hessian*

$$p''_{\lambda,0}(x)[h] := p''(x)[h] + \lambda p'(x)[h]^\top p'(x)[h].$$

which captures the conceptual idea. Suppose  $X \in (\mathbb{S}^{n \times n})^g$  and  $v \in \mathbb{R}^n$ . We say that the *modified Hessian is negative* at  $(X, v)$  if there is a  $\lambda_0 < 0$ , so that for all  $\lambda \leq \lambda_0$ ,

$$0 \leq -\langle p''_{\lambda,0}(X)[H]v, v \rangle$$

for all  $H \in (\mathbb{S}^{n \times n})^g$ . Given a subset  $\mathcal{R} = (\mathcal{R}(n))_{n=1}^\infty$ , with  $\mathcal{R}(n) \subseteq (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n$ , we say that the *modified Hessian is negative on  $\mathcal{R}$*  if it is negative at each  $(X, v) \in S$ .

Now we turn to motivation.

**Example 105.** *The classical  $n = 1$  case.* Suppose that  $p$  is strictly smoothly quasi-concave, meaning that all superlevel sets of  $p$  are strictly convex with strictly positively curved smooth boundary. Suppose that the gradient  $\nabla p$  (written as a row vector) never vanishes on  $\mathbb{R}^g$ . Then  $G = \nabla p(\nabla p)^\top$  is strictly positive, at each point  $X$  in  $\mathbb{R}^g$ . Fix such an  $X$ ; the modified Hessian can be decomposed as a block matrix subordinate to the tangent plane to the level set at  $X$ , denoted  $T_X$ , and to its orthogonal complement (the gradient direction):

$$T_X \oplus \{\lambda \nabla p : \lambda \in \mathbb{R}\}.$$

In this decomposition the modified Hessian has the form

$$R = \begin{bmatrix} A & B \\ B^\top & D + \lambda G \end{bmatrix}.$$

Here, in the case of  $\lambda = 0$ ,  $R$  is the Hessian and the second fundamental form is  $A$  or  $-A$ , depending on convention and the rather arbitrary choice of inward or outward normal to  $\nu$ . If we select our normal direction to be  $\nabla p$ , then  $-A$  is the classical second fundamental form as is consistent with the choice of sign in our definition in Subsection 6.1.3. (All this concern with the sign is unimportant to the content of this chapter and can be ignored by the reader.)

Next, in view of the presumed strict positive curvature of each level set  $\nu$ , the matrix  $A$  at each point of  $\nu$  is negative definite but the Hessian could have a negative eigenvalue. However, by standard Schur complement arguments,  $R$  will be negative definite if

$$D + \lambda G - B^\top A^{-1} B \prec 0$$

on this region. Thus, strict convexity assumptions on the sublevel sets of  $p$  make the modified Hessian negative definite for negative enough  $\lambda$ . One can make this negative definiteness uniform in  $X$  in various neighborhoods under modest assumptions.

Very unfortunately in the noncommutative case, Remark 6.8 [DHM11] implies that if  $n$  is large enough, then the second fundamental form will have a nonzero null space, thus strict negative definiteness of the  $A$  part of the modified Hessian is impossible.

Our trick, to deal with the likely reality that  $A$  is only positive semidefinite, and obtain a negative definite  $R$ , is to add another negative term, say  $\delta I$ , with arbitrarily small  $\delta < 0$ . After adding such  $\delta$ , the argument based on choosing  $-\lambda$  large succeeds as before. This  $\delta$  term plus the  $\lambda$  term produces the “relaxed Hessian”, to be introduced next, and proper selection of these terms make it negative definite.

**6.4.2. The relaxed Hessian.** Recall Let  $V_k(x)[h]$  denotes the vector of polynomials with entries  $h_j w(x)$ , where  $w \in \langle x \rangle$  runs through the set of  $g^k$  words of length  $k$ ,  $j = 1, \dots, g$ . Although the order of the entries is fixed in some of our earlier applications (see e.g. [DHM07b, (2.3)]) it is irrelevant for the moment. Thus,  $V_k = V_k(x)[h]$  is a vector of height  $g^{k+1}$ , and the vectors

$$V(x)[h] = \text{col}(V_0, \dots, V_{d-2}) \quad \text{and} \quad \tilde{V}(x)[h] = \text{col}(V_0, \dots, V_{d-1})$$

are vectors of height  $g\sigma(d-2)$  and  $g\sigma(d-1)$  respectively. Note that

$$\tilde{V}(x)[h]^\top \tilde{V}(x)[h] = \sum_{j=1}^g \sum_{\deg(w) \leq d-1} w(x)^\top h_j^2 w(x).$$

The *relaxed Hessian* of the symmetric nc polynomial  $p$  of degree  $d$  is defined to be

$$p''_{\lambda, \delta}(x)[h] := p''_{\lambda, 0}(x)[h] + \delta \tilde{V}(x)[h]^\top \tilde{V}(x)[h] \in \mathbb{R}\langle x \rangle[h].$$

Suppose  $X \in (\mathbb{S}^{n \times n})^g$  and  $v \in \mathbb{R}^n$ . We say that the *relaxed Hessian is negative* at  $(X, v)$  if for each  $\delta < 0$  there is a  $\lambda_\delta < 0$ , so that for all  $\lambda \leq \lambda_\delta$ ,

$$0 \leq -\langle p''_{\lambda, \delta}(X)[H]v, v \rangle$$

for all  $H \in (\mathbb{S}^{n \times n})^g$ . Given a  $\mathcal{R} = (\mathcal{R}(n))_{n=1}^\infty$ , with  $\mathcal{R}(n) \subseteq (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n$ , we say that the *relaxed Hessian is positive (resp., negative) on  $\mathcal{R}$*  if it is positive (resp., negative) at each  $(X, v) \in \mathcal{R}$ .

The following theorem provides a link between the signature of the clamped second fundamental form with that of the relaxed Hessian.



**Theorem 106.** *Suppose  $p$  is a symmetric nc polynomial of degree  $d$  in  $g$  symmetric variables and  $(X, v) \in (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n$ . If  $\mathcal{V}(p)$  has positive curvature at  $(X, v) \in \mathcal{V}_n(p)$ , i.e., if*

$$\langle p''(X)[H]v, v \rangle \leq 0 \quad \text{for every } H \in \mathcal{T}_p(X, v),$$

*then for every  $\delta < 0$  there exists a  $\lambda_\delta < 0$  such that for all  $\lambda \leq \lambda_\delta$ ,*

$$\langle p''_{\lambda, \delta}(X)[H]v, v \rangle \leq 0 \quad \text{for every } H \in (\mathbb{S}^{n \times n})^g;$$

*i.e., the relaxed Hessian of  $p$  is negative at  $(X, v)$ .*

We leave the proof of Theorem 106 to the reader.

The basic idea of the proof of Theorem 103, is to obtain a negative relaxed Hessian  $q$  from Theorem 106 and then apply Theorem 104. We begin with the following lemma.

**Lemma 107.** *Suppose  $R$  and  $T$  are operators on a finite dimensional Hilbert space  $H = K \oplus L$ . Suppose further that, with respect to this decomposition of  $H$ , the operator  $R = CC^\top$  for*

$$C = \begin{bmatrix} r \\ c \end{bmatrix} : L \rightarrow K \oplus L \quad \text{and} \quad T = \begin{bmatrix} T_0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*If  $c$  is invertible and if for every  $\delta > 0$  there is a  $\eta > 0$  such that for all  $\lambda > \eta$ ,*

$$T + \delta I + \lambda R \succeq 0,$$

*then  $T \succeq 0$ .*

*Proof.* Write

$$T + \delta I + \lambda R = \begin{bmatrix} T_0 + \delta I + \lambda r r^\top & \lambda r c^\top \\ \lambda c r^\top & \delta + \lambda c c^\top \end{bmatrix}.$$

From Schur complements it follows that

$$T_0 + \delta I + r(\lambda - \lambda^2 c^\top (\delta + \lambda c c^\top)^{-1} c) r^\top \succeq 0.$$

Now

$$\begin{aligned} r(\lambda - \lambda^2 c^\top (\delta + \lambda c c^\top)^{-1} c) r^\top &= \lambda r c^\top ((c c^\top)^{-1} - \lambda (\delta + \lambda c c^\top)^{-1}) c r^\top \\ &= \lambda r c^\top \delta (c c^\top)^{-1} (\delta + \lambda (c c^\top))^{-1} c r^\top \\ &\preceq \delta r (c c^\top)^{-1} r^\top. \end{aligned}$$

Hence,

$$T_0 + \delta I + \delta r (c c^\top)^{-1} r^\top \succeq 0.$$

Since the above inequality holds for all  $\delta > 0$ , it follows that  $T_0 \succeq 0$ . ■

We now have enough machinery developed to prove Theorem 103.

*Proof of Theorem 103.* Fix  $\lambda, \delta > 0$  and consider  $q(x)[h] = -p''_{\lambda, \delta}(x)[h]$ . We are led to investigate the middle matrix  $Z^{\lambda, \delta}$  of  $q(x)[h]$ , whose border vector  $V(x)[h]$  includes all monomials of the form  $h_j m$ , where  $m$  is a word in  $x$  only of length at most  $d - 1$ ; here  $d$  is the degree of  $p$ . Indeed,

$$Z^{\lambda, \delta} = Z + \delta I + \lambda W,$$

where  $Z$  is the middle matrix for  $-p''(x)[h]$ , and  $W$  is the middle matrix for the polynomial  $p'(x)[h]^\top p'(x)[h]$ . With an appropriate choice of ordering for the border vector  $V$ , we have,  $W = CC^\top$ , where

$$C(x) = \begin{bmatrix} w(x) \\ c \end{bmatrix},$$

for a nonzero vector  $c$ ; and at the same time,

$$Z(x) = \begin{bmatrix} Z^{0,0}(x) & 0 \\ 0 & 0 \end{bmatrix}.$$

By the curvature hypothesis at a given  $X$  with  $(X, v) \in \mathcal{R}$ , Theorem 106 implies for every  $\delta > 0$  there is an  $\eta > 0$  such that if  $\lambda > \eta$

$$\langle q(X)[H]v, v \rangle \geq 0 \quad \text{for all } (X, v) \in \mathcal{R} \text{ and all } H.$$

Hence, by Theorem 104, the middle matrix,  $Z^{\lambda, \delta}(X)$  for  $q(x)[h]$  is positive semidefinite. We are in the setting of Lemma 107 from which we obtain  $Z^{0,0}(X) \succeq 0$ . If this held for  $X$  in a nc basic open semialgebraic set, then Theorem 71 forces  $p$  to have degree no greater than 2. The proof of that theorem applies easily here to finish this proof. ■

## 6.5. Exercises.

**Exercise 108.** Compute the BV-MM representation for the relaxed Hessian of  $x^3$  and  $x^4$ .

## 7. CONVEX SEMIALGEBRAIC NC SETS

In this section we will give a brief overview of convex semialgebraic nc sets and positivity of nc polynomials on them. We shall see that their structure is much more rigid than that of their commutative counterparts. For example, roughly speaking, each convex semialgebraic nc set is a spectrahedron; i.e., a solution set of a linear

matrix inequality (cf. Subsection 7.1 below). Similarly, every nc polynomial nonnegative on a spectrahedron admits a sum of squares representation with weights and optimal degree bounds (see Subsection 7.2 for details and precise statements).

**7.1. nc Spectrahedra.** Let  $L$  be an affine linear pencil. Then the solution set of the linear matrix inequality (LMI)  $L(x) \succ 0$  is

$$\mathcal{D}_L = \bigcup_{n \in \mathbb{N}} \{X \in (\mathbb{S}^{n \times n})^g : L(X) \succ 0\},$$

and is called a *nc spectrahedron*. The set  $\mathcal{D}_L$  is convex in the sense that each

$$\mathcal{D}_L(n) := \{X \in (\mathbb{S}^{n \times n})^g : L(X) \succ 0\}$$

is convex. It is also a noncommutative basic open semialgebraic set as defined in Subsection 2.1.2 above. The main theorem of this section is the converse, a result which has implications for both semidefinite programming and systems engineering.

Most of the time we will focus on monic linear pencils. An affine linear pencil  $L$  is called *monic* if  $L(0) = I$ , i.e.,  $L(x) = I + A_1 x_1 + \cdots + A_g x_g$ . Since we are mostly interested in the set  $\mathcal{D}_L$ , there is no harm in reducing to this case whenever  $\mathcal{D}_L \neq \emptyset$ ; see Exercise 111.

Let  $p \in \mathbb{R}^{\delta \times \delta} \langle x \rangle$  be a given symmetric noncommutative  $\delta \times \delta$ -valued matrix polynomial. Assuming that  $p(0) \succ 0$ , the positivity set  $\mathcal{D}_p(n)$  of a noncommutative symmetric polynomial  $p$  in dimension  $n$  is the component of 0 of the set

$$\{X \in (\mathbb{S}^{n \times n})^g : p(X) \succ 0\}.$$

The *positivity set*,  $\mathcal{D}_p$ , is the sequence of sets  $(\mathcal{D}_p(n))_{n \in \mathbb{N}}$ . The noncommutative set  $\mathcal{D}_p$  is called *convex* if, for each  $n$ ,  $\mathcal{D}_p(n)$  is convex.

**Theorem 109** (Helton-McCullough [HM12]). *Fix  $p$  a  $\delta \times \delta$  symmetric matrix of polynomials in noncommuting variables. Assume*

- (1)  $p(0)$  is positive definite;
- (2)  $\mathcal{D}_p$  is bounded; and
- (3)  $\mathcal{D}_p$  is convex.

*Then there is a monic linear pencil  $L$  such that*

$$\mathcal{D}_L = \mathcal{D}_p.$$

Here we shall confine ourselves to a few words about the techniques involved in the proof, and refer the reader to [HM12] for the full proof. Since we are dealing with matrix convex sets, it is not surprising that the starting point for our analysis is

the matricial version of the Hahn-Banach Separation theorem of Effros and Winkler [EW97] which (itself a part of the theory of operator spaces and completely positive maps [BL04, Pau02, Pis03]) says that given a point  $x$  not inside a matrix convex set there is a (finite) linear matrix inequality which separates  $x$  from the set. For a general matrix convex set  $\mathcal{C}$ , the conclusion is then that there is a collection, likely infinite, of LMIs which cut out  $\mathcal{C}$ .

In the case  $\mathcal{C}$  is matrix convex and also semialgebraic, the challenge is to prove that there is actually a *finite* collection of LMIs which define  $\mathcal{C}$ . The techniques used to meet this challenge have little relation to the methods of noncommutative calculus and positivity in the previous sections. Indeed a basic tool (of independent interest) is a degree bounded type of free Zariski closure of a single point  $(X, v) \in (\mathbb{S}^{n \times n})^g \times \mathbb{R}^n$ ,

$$Z_d(X, v) := \bigcup_m \{(Y, w) \in (\mathbb{S}^{m \times m})^g \times \mathbb{R}^m : q(Y)w = 0 \text{ if } q(X)v = 0, q \in \mathbb{R}\langle x \rangle_d\}.$$

Chief among a pleasant list of natural properties is the fact that there is an  $(X, v)$  with  $X \in \partial\mathcal{D}_p$  and  $p(X)v = 0$  for which  $Z_d(X, v)$  contains all pairs  $(Y, w)$  such that  $Y \in \partial\mathcal{D}_p$  and  $p(Y)w = 0$ . Combining this with the Effros-Winkler Theorem and battling degeneracies is a bit tricky, but voila separation prevails in the end. See [HM12] for the details.

An unexpected consequence of Theorem 109 is that projections of noncommutative semialgebraic sets may not be semialgebraic, see Exercise 112. For perspective, in the commutative case of a basic open semialgebraic subset  $\mathcal{C}$  of  $\mathbb{R}^g$ , there is a stringent condition, called the “*line test*” (see Chapter 6 for more details), which, in addition to convexity, is necessary for  $\mathcal{C}$  to be a spectrahedron. In two dimensions the line test is necessary and sufficient [HV07], a result used by Lewis-Parrilo-Ramana [LPR05] to settle a 1958 conjecture of Peter Lax on hyperbolic polynomials.

In summary, if a (commutative) bounded basic open semialgebraic convex set is a spectrahedron, then it must pass the highly restrictive line test; whereas a nc basic open semialgebraic set is a spectrahedron if and only if it is convex.

**7.2. Noncommutative Positivstellensätze under convexity assumptions.** An algebraic certificate for positivity of a polynomial  $p$  on a semialgebraic set  $S$  is a Positivstellensatz. The familiar fact that a polynomial  $p$  in one-variable which is positive on  $\mathbb{R}$  is a sum of squares is an example.

The theory of Positivstellensätze - a pillar of the field of real algebraic geometry - underlies the main approach currently used for global optimization of polynomials. See [Las10] or Chapters 2 and 3 of Parrilo for a beautiful treatment of this, and other,

applications of commutative real algebraic geometry. Further, because convexity of a polynomial  $p$  on a set  $S$  is equivalent to positivity of the Hessian of  $p$  on  $S$ , this theory also provides a link between convexity and semialgebraic geometry. Indeed, this link in the noncommutative setting ultimately lead to the conclusion the a matrix convex noncommutative polynomial has degree at most two, cf. Section 4.4.

In this section we give a result of opposite type. We present a noncommutative Positivstellensatz for a polynomial to be nonnegative on a convex semialgebraic nc set (i.e., on a spectrahedron). Again, this result is cleaner and more rigid than the commutative counterparts (cf. Theorem 10).

**Theorem 110** ([HKM12a]). *Suppose  $L$  is a monic linear pencil. Then a noncommutative polynomial  $p$  is positive semidefinite on  $\mathcal{D}_L$  if and only if it has a weighted sum of squares representation with optimal degree bounds. Namely,*

$$p = s^\top s + \sum_j^{\text{finite}} f_j^\top L f_j, \quad (26)$$

where  $s, f_j$  are vectors of noncommutative polynomials of degree no greater than  $\frac{\deg(p)}{2}$ .

The main ingredient of the proof is an analysis of rank preserving extensions of truncated noncommutative Hankel matrices; see [HKM12a] for details. We point out that with  $L = 1$ , Theorem 110 recovers Theorem 10.

Theorem 110 contrasts sharply with the commutative setting, where the degrees of  $s, f_j$  are vastly greater than  $\deg(p)$  and assuming only  $p$  nonnegative yields a clean Positivstellensatz so seldom that the cases are noteworthy.

### 7.3. Exercises.

**Exercise 111.** Suppose  $L$  is an affine linear pencil such that  $0 \in \mathcal{D}_L(1)$ . Show that there is a monic linear pencil  $\check{L}$  with  $\mathcal{D}_L = \mathcal{D}_{\check{L}}$ .

**Exercise 112.** Chapters 6 and 7 discuss sets  $D \subseteq \mathbb{R}^g$  which have a semidefinite representation as a strict generalization of a spectrahedron. For instance, consider the TV screen (cf. Subsection 2.1.2)

$$\text{ncTV}(1) = \{X \in \mathbb{R}^2: 1 - X_1^4 - X_2^4 > 0\} \subseteq \mathbb{R}^2.$$

Given  $\alpha$  a positive real number, choose  $\gamma^4 = 1 + 2\alpha^2$  and let

$$L_0 = \begin{bmatrix} 1 & 0 & y_1 \\ 0 & 1 & y_2 \\ y_1 & y_2 & 1 - 2\alpha(y_1 + y_2) \end{bmatrix} \quad (27)$$

and

$$L_j = \begin{bmatrix} 1 & \gamma x_j \\ \gamma x_j & \alpha + y_j \end{bmatrix}, \quad j = 1, 2. \quad (28)$$

Note that the  $L_j$  are not monic, but because  $L_j(0) \succ 0$ , they can be normalized to be monic without altering the solution sets of  $L_j(X) \succ 0$ , cf. Exercise 111. Let  $L = L_0 \oplus L_1 \oplus L_2$ .

It is readily verified that  $\text{ncTV}(1)$  is the projection, onto the first two (the  $x$ ) coordinates of the set  $\mathcal{D}_L(1)$ ; i.e.,

$$\text{ncTV}(1) = \{X \in \mathbb{R}^2 : \exists Y \in \mathbb{R}^2 \ L(X, Y) \succ 0\}.$$

- (1) Show that  $\text{ncTV}(1)$  is not a spectrahedron. (*Hint*: How often is  $L_{\text{TV}}(tX, tY)$  for  $t \in \mathbb{R}$  singular?)
- (2) Show that  $\text{ncTV}$  is not the projection of the nc spectrahedron  $\mathcal{D}_L$ .
- (3) Show that  $\text{ncTV}$  is not the projection of *any* nc spectrahedron.
- (4) Is  $\text{ncTV}(2)$  a projection of a spectrahedron? (Feel free to use the results about  $\text{ncTV}$  and LMI representable sets (spectrahedra), stated without proofs, from Subsection 2.1.2 and Subsection 7.1.)

**Exercise 113.** If  $q$  is a symmetric concave matrix-valued polynomial with  $q(0) = I$ , then there exists a linear pencil  $L$  and a matrix-valued linear polynomial  $\Lambda$  such that

$$q = I - L - \Lambda^\top \Lambda.$$

**Exercise 114.** Consider the monic linear pencil

$$M(x) = \begin{bmatrix} 1 & x \\ x & 1 \end{bmatrix}.$$

- (1) Determine  $\mathcal{D}_M$ .
- (2) Show that  $1 + x$  is positive semidefinite on  $\mathcal{D}_M$ .
- (3) Construct a representation for  $1 + x$  of the form (26).

**Exercise 115.** Consider the univariate affine linear pencil

$$L(x) = \begin{bmatrix} 1 & x \\ x & 0 \end{bmatrix}.$$

- (1) Determine  $\mathcal{D}_L$ .
- (2) Show that  $x$  is positive semidefinite on  $\mathcal{D}_L$ .
- (3) Does  $x$  admit a representation of the form (26)?

**Exercise 116.** Let  $L$  be an affine linear pencil. Prove that:

- (1)  $\mathcal{D}_L$  is bounded if and only if  $\mathcal{D}_L(1)$  is bounded;
- (2)  $\mathcal{D}_L = \emptyset$  if and only if  $\mathcal{D}_L(1) = \emptyset$ .

**Exercise 117.** Let  $L = I + A_1x_1 + \cdots + A_gx_g$  be a monic linear pencil and assume that  $\mathcal{D}_L(1)$  is bounded. Show that  $I, A_1, \dots, A_g$  are linearly independent.

**Exercise 118.** Let

$$\Delta(x_1, x_2) = I + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} x_2 = \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & 0 \\ x_2 & 0 & 1 \end{bmatrix}$$

and

$$\Gamma(x_1, x_2) = I + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} x_1 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_2 = \begin{bmatrix} 1+x_1 & x_2 \\ x_2 & 1-x_1 \end{bmatrix}$$

be affine linear pencils. Show:

- (1)  $\mathcal{D}_\Delta(1) = \mathcal{D}_\Gamma(1)$ .
- (2)  $\mathcal{D}_\Gamma(2) \subsetneq \mathcal{D}_\Delta(2)$ .
- (3) Is  $\mathcal{D}_\Delta \subseteq \mathcal{D}_\Gamma$ ? What about  $\mathcal{D}_\Gamma \subseteq \mathcal{D}_\Delta$ ?

**Exercise 119.** Let  $L = A_1x_1 + \cdots + A_gx_g \in \mathbb{S}^{d \times d} \langle x \rangle$  be a (homogeneous) linear pencil. Then the following are equivalent:

- (i)  $\mathcal{D}_L(1) \neq \emptyset$ ;
- (ii) If  $u_1, \dots, u_m \in \mathbb{R}^d$  with  $\sum_{i=1}^m u_i^\top L(x) u_i = 0$ , then  $u_1 = \cdots = u_m = 0$ .

## 8. FROM FREE REAL ALGEBRAIC GEOMETRY TO THE REAL WORLD

Now that you have gone through the mathematics we return to its implications. In the linear systems engineering problems you have seen both in Subsection 1.1 and in Chapter 2.2.1, the conclusion was that the problem was equivalent to solving an LMI. Indeed this is what one sees throughout the literature. Thousands of engineering papers have a dimension free problem and it converts (often by serious cleverness) to an LMI in the best of cases, or more likely there is some approximate solution which is an LMI.

While engineers would be satisfied with convexity, what they actually do get is an LMI. One would hope that there is a rich world of convex situations not equivalent to an LMI. Then there would be a variety of methods waiting to be discovered for dealing with them. Alas what we have shown here is compelling evidence that any convex dimension free problem is equivalent to an LMI. Thus there is no rich world of convexity beyond what is already known and no armada of techniques beyond those for producing LMIs which we already see all around us.

## REFERENCES

- [BM+] S. Balasubramanian, S. McCullough: Quasi-convex free polynomials, to appear in *Proc. Amer. Math. Soc.*, <http://arxiv.org/abs/1208.3582> 8
- [Ber76] G.M. Bergman: Rational relations and rational identities in division rings I, *J. Algebra* **43** (1976) 252–266 20
- [Bha97] R. Bhatia: *Matrix Analysis*, Springer-Verlag, 1997 42
- [BL04] D.P. Blecher, C. Le Merdy: *Operator algebras and their modules—an operator space approach*, Oxford Science Publications, 2004 60
- [BCR98] J. Bochnak, M. Coste, M.F. Roy: *Real algebraic geometry*, Springer-Verlag, 1998 51
- [BK13] M. Brešar, I. Klep: A local-global principle for linear dependence of noncommutative polynomials, *Israel J. Math.* **193** (2013) 71–82 49
- [CKP10] K. Cafuta, I. Klep, J. Povh: A note on the nonexistence of sum of squares certificates for the Bessis-Moussa-Villani conjecture, *J. math. phys.* **51** (2010) 083521 8
- [CKP11] K. Cafuta, I. Klep, J. Povh: **NCSOStools**: a computer algebra system for symbolic and numerical computation with noncommutative polynomials, *Optim. Methods Softw.* **26** (2011) 363–380 30
- [CHSY03] J.F. Camino, J.W. Helton, R.E. Skelton, J. Ye: Matrix inequalities: A symbolic procedure to determine convexity automatically, *Integral Equations and Operator Theory* **46** (2003) 399–454 30, 34, 42, 43, 44
- [Coh95] P.M. Cohn: *Skew fields. Theory of general division rings*, Cambridge University Press, 1995 19, 23, 27
- [Coh06] P.M. Cohn: *Free ideal rings and localization in general rings*, Cambridge University Press, 2006 19, 27
- [DLTW08] A.C. Doherty, Y.-C. Liang, B. Toner, S. Wehner: The quantum moment problem and bounds on entangled multi-prover games. In *Twenty-Third Annual IEEE Conference on Computational Complexity* (2008) 199–210 8
- [dOHMP09] M. de Oliveira, J.W. Helton, S. McCullough, M. Putinar: Engineering Systems and Free Semi-Algebraic Geometry, In: *Emerging Applications of Algebraic Geometry*, 17–62, IMA Vol. Math. Appl. **149**, Springer-Verlag, 2009 2, 3, 9
- [DGHM09] H. Dym, J.M. Greene, J.W. Helton, S. McCullough: Classification of all noncommutative polynomials whose Hessian has negative signature one and a noncommutative second fundamental form, *J. Anal. Math.* **108** (2009) 19–59 39, 50
- [DHM07a] H. Dym, J.W. Helton, S. McCullough: Irreducible noncommutative defining polynomials for convex sets have degree four or less, *Indiana Univ. Math. J.* **56** (2007) 1189–1232 13, 52
- [DHM07b] H. Dym, J.W. Helton, S. McCullough: The Hessian of a Non-commutative Polynomial has Numerous Negative Eigenvalues, *J. Anal. Math.* **102** (2007) 29–76 50, 56
- [DHM11] H. Dym, J.W. Helton, S. McCullough: Noncommutative varieties with curvature having bounded signature, *Illinois J. Math.* **55** (2011) 427–464 56
- [Eff09] E.G. Effros: A matrix convexity approach to some celebrated quantum inequalities, *Proc. Natl. Acad. Sci. USA* **106** (2009) 1006–1008 8
- [ENG11] A. Ebadiana, I. Nikoufarb, M.E. Gordjic: Perspectives of matrix convex functions, *Proc. Natl. Acad. Sci. USA* **108** (2011) 7313–7314 9



- [EW97] E.G. Effros, S. Winkler: Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems, *J. Funct. Anal.* **144** (1997) 117–152 [60](#)
- [GL05] W.S. Gray, Y. Li: Generating Series for Interconnected Analytic Nonlinear Systems, *SIAM J. Control Optim.* **44** (2005) 646–672 [9](#)
- [GT12] W.S. Gray, M. Thitsa: A Unified Approach to Generating Series for Mixed Cascades of Analytic Nonlinear Input-Output Systems, *International J. Control* **85** (2012) 1737–1754 [9](#)
- [GHV11] J.M. Greene, J.W. Helton, V. Vinnikov: Noncommutative Plurisubharmonic Polynomials Part I: Global Assumptions, *J. Funct. Anal.* **261** (2011) 3390–3417 [27](#)
- [Han97] F. Hansen: Operator convex functions of several variables, *Publ. Res. Inst. Math. Sci.* **33** (1997) 443–463 [42](#)
- [BT07] F. Hansen, J. Tomiyama: Differential analysis of matrix convex functions, *Linear Algebra Appl.* **420** (2007) 102–116 [42](#)
- [HHLM08] D.M. Hay, J.W. Helton, A. Lim, S. McCullough: Non-commutative partial matrix convexity, *Indiana Univ. Math. J.* **57** (2008) 2815–2842 [8](#)
- [Hel02] J.W. Helton: “Positive” noncommutative polynomials are sums of squares, *Ann. of Math. (2)* **156** (2002) 675–694 [2](#), [16](#)
- [HKM10a] J.W. Helton, I. Klep, S. McCullough: Analytic mappings between noncommutative pencil balls, *J. Math. Anal. Appl.* **376** (2011) 407–428 [2](#), [19](#)
- [HKM10b] J.W. Helton, I. Klep, S. McCullough: Proper Analytic Free Maps, *J. Funct. Anal.* **260** (2011) 1476–1490 [2](#), [19](#)
- [HKM10c] J.W. Helton, I. Klep, S. McCullough: Relaxing LMI Domination Matricially, In *49th IEEE Conference on Decision and Control* (2010) 3331–3336 [2](#)
- [HKM12a] J.W. Helton, I. Klep, S. McCullough: The convex Positivstellensatz in a free algebra, *Adv. Math.* **231** (2012) 516–534 [2](#), [61](#)
- [HKM12b] J.W. Helton, I. Klep, S. McCullough: Convexity and Semidefinite Programming in dimension-free matrix unknowns, In: *Handbook of Semidefinite, Cone and Polynomial Optimization* edited by M. Anjos and J. B. Lasserre, 377–405, Springer-Verlag, 2012 [9](#)
- [HKM13] J.W. Helton, I. Klep, S. McCullough: The matricial relaxation of a linear matrix inequality, *Math. Program.* **138** (2013) 401–445 [2](#)
- [HKMS09] J.W. Helton, I. Klep, S. McCullough, N. Slindend: Noncommutative ball maps, *J. Funct. Anal.* **257** (2009) 47–87 [2](#)
- [HMe98] J.W. Helton, O. Merino: Sufficient conditions for optimization of matrix functions. In *37th IEEE Conference on Decision and Control* (1998) 3361–3365 [39](#)
- [HM04a] J.W. Helton, S. McCullough: Convex noncommutative polynomials have degree two or less, *SIAM J. Matrix Anal. Appl.* **25** (2004) 1124–1139 [39](#), [42](#)
- [HM04b] J.W. Helton, S. McCullough: A Positivstellensatz for noncommutative polynomials, *Trans. Amer. Math. Soc.* **356** (2004) 3721–3737 [8](#)
- [HM12] J.W. Helton, S. McCullough: Every free basic convex semialgebraic set has an LMI representation, *Ann. of Math. (2)* **176** (2012) 979–1013 [2](#), [13](#), [59](#), [60](#)
- [HMP04] J.W. Helton, S. McCullough, M. Putinar: A non-commutative Positivstellensatz on isometries, *J. reine angew. Math.* **568** (2004) 71–80 [2](#)
- [HMPV09] J.W. Helton, S. McCullough, M. Putinar, V. Vinnikov: Convex matrix inequalities versus linear matrix inequalities, *IEEE Trans. Automat. Control* **54** (2009) 952–964 [9](#)

- [HMV06] J.W. Helton, S. McCullough, V. Vinnikov: Noncommutative convexity arises from linear matrix inequalities, *J. Funct. Anal.* **240** (2006) 105–191 [19](#), [23](#)
- [HOMS+] J.W. Helton, M. de Oliveira, R.L. Miller, M. Stankus: [NCAlgebra](#): A Mathematica package for doing non commuting algebra, available from <http://www.math.ucsd.edu/~ncalg/> [28](#)
- [HP07] J.W. Helton, M. Putinar: Positive Polynomials in Scalar and Matrix Variables, the Spectral Theorem and Optimization, In: *Operator theory, structured matrices, and dilations*, Theta Ser. Adv. Math., **7** (2007) 229–306 [9](#)
- [HV07] J.W. Helton, V. Vinnikov: Linear matrix inequality representation of sets, *Comm. Pure Appl. Math.* **60** (2007) 654–674 [13](#), [60](#)
- [KVV09] D. Kalyuzhnyi-Verbovetskiĭ, V. Vinnikov: Singularities of rational functions and minimal factorizations: the noncommutative and the commutative setting, *Linear Algebra Appl.* **430** (2009) 869–889 [19](#), [23](#)
- [KVV+] D. Kalyuzhnyi-Verbovetskiĭ, V. Vinnikov: Foundations of noncommutative function theory, *preprint*, <http://arxiv.org/abs/1212.6345> [2](#), [19](#)
- [KS08a] I. Klep, M. Schweighofer: Connes’ embedding conjecture and sums of Hermitian squares, *Adv. Math.* **217** (2008) 1816–1837 [2](#), [8](#)
- [KS08b] I. Klep, M. Schweighofer: Sums of Hermitian squares and the BMV conjecture, *J. Stat. Phys* **133** (2008) 739–760 [2](#), [8](#)
- [Kra36] F. Kraus: Über konvexe Matrixfunktionen, *Math. Z.* **41** (1936) 18–42 [42](#)
- [Las10] J.B. Lasserre: *Moments, positive polynomials and their applications*, Imperial College Press, 2010 [60](#)
- [LPR05] A.S. Lewis, P.A. Parrilo, M.V. Ramana: The Lax conjecture is true, *Proc. Amer. Math. Soc.* **133** (2005) 2495–2499 [60](#)
- [LCL04] T. Lyons, M. Caruana, T. Lévy: Differential equations driven by rough paths, in *École d’Eté de Probabilités de Saint-Flour XXXIV*, Lecture Notes in Mathematics 1908, Springer-Verlag, Berlin, 2004 [9](#)
- [McC01] S. McCullough: Factorization of operator-valued polynomials in several noncommuting variables, *Linear Algebra Appl.* **326** (2001) 193–203 [2](#), [16](#)
- [MS11] P.S. Muhly, B. Solel: Progress in noncommutative function theory, *Sci. China Ser. A* **54** (2011) 2275–2294 [2](#)
- [NFBK10] B. Sz.-Nagy, C. Foias, H. Bercovici, L. Kerchy: *Harmonic analysis of Operators on Hilbert Space*, Springer-Verlag, 2010 [26](#)
- [OST07] H. Osaka, S. Silvestrov, J. Tomiyama: Monotone operator functions, gaps and power moment problem, *Math. Scand.* **100** (2007) 161–183 [42](#)
- [Pau02] V. Paulsen: *Completely bounded maps and operator algebras*, Cambridge University Press, 2002 [60](#)
- [Pis03] G. Pisier: *Introduction to operator space theory*, Cambridge University Press, 2003 [60](#)
- [PNA10] S. Pironio, M. Navascués, A. Acín: Convergent relaxations of polynomial optimization problems with noncommuting variables, *SIAM J. Optim.*, **20** (2010) 2157–2180 [2](#), [8](#)
- [Pop06] G. Popescu: Free holomorphic functions on the unit ball of  $\mathcal{B}(\mathcal{H})^n$ , *J. Funct. Anal.* **241** (2006) 268–333 [2](#), [19](#)
- [Pop10] G. Popescu: Free holomorphic automorphisms of the unit ball of  $B(H)^n$ , *J. reine angew. Math.* **638** (2010) 119–168 [19](#)

- [Smü05] K. Schmüdgen: A strict Positivstellensatz for the Weyl algebra, *Math. Ann.* **331** (2005) 779–794 [2](#), [9](#)
- [Smü09] K. Schmüdgen: Noncommutative real algebraic geometry - some basic concepts and first ideas, In: *Emerging applications of algebraic geometry*, 325-350, IMA Vol. Math. Appl. **149**, Springer-Verlag, 2009 [2](#), [9](#)
- [SV06] D. Shlyakhtenko, D.-V. Voiculescu: Free analysis workshop summary: American institute of mathematics,  
<http://www.aimath.org/pastworkshops/freeanalysis.html> [2](#), [9](#), [19](#)
- [SI95] R.E. Skelton, T. Iwasaki: Eye on Education: Increased Roles of Linear Algebra in Control Education, *IEEE Control Syst. Mag.* **15** (1995) 76-90 [9](#)
- [SIG97] R.E. Skelton, T. Iwasaki, K.M. Grigoriadis: *A Unified Algebraic Approach to Linear Control Design*, Taylor & Francis, 1997 [9](#)
- [Tay73] J.L. Taylor: Functions of several noncommuting variables, *Bull. Amer. Math. Soc.* **79** (1973) 1–34 [19](#)
- [Uch02] M. Uchiyama: Operator monotone functions and operator inequalities, *Sugaku Expositions* **18** (2005) 39–52 [42](#)
- [Voi04] D.-V. Voiculescu: Free analysis questions I: Duality transform for the coalgebra of  $\partial_{X:B}$ , *International Math. Res. Notices* **16** (2004) 793–822 [2](#), [19](#)
- [Voi10] D.-V. Voiculescu: Free analysis questions II: The Grassmannian completion and the series expansions at the origin, *J. reine angew. Math.* **645** (2010) 155–236 [2](#), [19](#)
- [VDN92] D.-V. Voiculescu, K.J. Dykema, A. Nica: *Free random variables. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups*, Amer. Math. Soc., 1992 [9](#)

J. WILLIAM HELTON, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO

*E-mail address:* `helton@math.ucsd.edu`

IGOR KLEP, DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF AUCKLAND, NEW ZEALAND

*E-mail address:* `igor.klep@auckland.ac.nz`

SCOTT McCULLOUGH, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE

*E-mail address:* `sam@math.ufl.edu`

## NOT FOR PUBLICATION

## CONTENTS

1. Introduction .....	1
1.1. Motivation .....	2
1.1.1. Linear Systems Engineering .....	2
1.1.2. Linear systems .....	3
1.1.3. Connecting linear systems .....	3
1.1.4. Energy dissipation .....	5
1.1.5. Schur Complements and Linear Matrix Inequalities .....	6
1.1.6. Putting it together .....	7
1.1.7. Quantum Phenomena .....	8
1.1.8. Miscellaneous applications .....	9
1.2. Further reading .....	9
1.3. Guide to the chapter .....	9
2. Basics of nc Polynomials and their Convexity .....	10
2.1. Noncommutative polynomials .....	10
2.1.1. Noncommutative convexity for polynomials .....	11
2.1.2. Noncommutative polynomial inequalities and convexity .....	12
2.2. Noncommutative polynomials, the formalities .....	13
2.2.1. Noncommutative matrix polynomials .....	14
2.2.2. Linear pencils .....	14
2.2.3. Polynomial evaluations .....	15
2.3. Noncommutative convexity revisited and nc positivity .....	16
2.4. Directional derivatives vs. nc convexity and positivity .....	17
2.5. Symmetric, free, mixed, and classes of variables .....	18
2.5.1. Free variables .....	18
2.5.2. Mixed variables .....	19
2.6. Noncommutative rational functions .....	19
2.6.1. Rational functions, a gentle introduction .....	20
2.6.2. Matrices of Rational Functions; $LDL^\top$ .....	21

2.6.3. More on rational functions.....	23
2.7. Exercises .....	23
3. Computer algebra support.....	28
3.1. NCAgebra .....	28
3.1.1. Warning.....	30
3.2. NCSOSTools.....	30
4. A Gram-like representation .....	32
4.1. Illustrating the ideas.....	32
4.1.1. The positivity of $q$ vs. positivity of the MM.....	34
4.2. Details of the Middle Matrix representation.....	36
4.3. The Middle Matrix of $p''$ .....	38
4.4. Positivity of the Middle Matrix and the demise of nc convexity .....	39
4.5. The signature of the middle matrix .....	40
4.6. Exercises .....	41
4.7. A glimpse of history .....	42
5. Der QuadratischePositivstellensatz.....	42
5.1. The Camino, Helton, Skelton, Ye (CHSY) Lemma .....	43
5.2. Linear Dependence of Symbolic Functions .....	44
5.2.1. Direct Sums .....	46
5.2.2. Main Result on Linear Dependence .....	47
5.2.3. Proof of Theorem 92.....	47
5.3. Proof of the QuadratischePositivstellensatz .....	48
5.4. Exercises .....	49
6. NC varieties with positive curvature have degree two.....	49
6.1. NC varieties and their curvature.....	50
6.1.1. Varieties, tangent planes, and the second fundamental form.....	50
6.1.2. Full rank points .....	51
6.1.3. Positive curvature .....	51
6.1.4. Irreducibility: The minimum degree defining polynomial condition .....	52
6.2. A very simple example.....	52
6.3. Main Result: Positive curvature and the degree of $p$ .....	53

6.4. Ideas and proofs .....	54
6.4.1. The modified Hessian .....	54
6.4.2. The relaxed Hessian .....	56
6.5. Exercises .....	58
7. Convex semialgebraic nc sets .....	58
7.1. nc Spectrahedra .....	59
7.2. Noncommutative Positivstellensätze under convexity assumptions .....	60
7.3. Exercises .....	61
8. From free real algebraic geometry to the real world .....	63
References .....	64